Numerical Approximations for Non-Zero-Sum Stochastic Differential Games

Harold J.Kushner*
Applied Math. Dept., Brown University
Providence RI 02912
hjk@dam.brown.edu

December 10, 2005

Running Head: Numerical Approximations for Games

Abstract

The Markov chain approximation method is a widely used, and efficient family of methods for the numerical solution a large part of stochastic control problems in continuous time for reflected-jump-diffusion-type models. It converges under broad conditions, and there are good algorithms for solving the numerical approximations if the dimension is not too high. It has been extended to zero-sum stochastic differential games. We apply the method to consider a class of non-zero stochastic differential games with a diffusion system model where the controls for the two players are separated in the dynamics and cost function. There have been successful applications of the algorithms, but convergence proofs have been lacking. It is shown that equilibrium values for the approximating chain converge to equilibrium values for the original process and that any equilibrium value for the original process can be approximated by an ϵ-equilibrium for the chain for arbitrarily small $\epsilon > 0$. The numerical method solves a stochastic game for a finite-state Markov chain.

Keywords: stochastic differential games, non-zero-sum games, numerical methods, Markov chain approximations
2000 Mathematics Subject classifications: 60F17, 65C30, 65C40, 91A15, 91A23, 93E25

1 Introduction

The aim of this paper is to extend the Markov chain approximation method to numerically solve non-zero-sum stochastic differential games. The method is widely used, robust, and relatively easy to use. It covers the majority of

*This work was partially supported by NSF grant DMS-0506928 and ARO contract W911NF-05-10928

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **10 DEC 2005**	2. REPORT TYPE	3. DATES COVERED **00-12-2005 to 00-12-2005**

4. TITLE AND SUBTITLE **Numerical Approximations for Non-Zero-Sum Stochastic Differential Games**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Brown University,Division of Applied Mathematics,182 George Street,Providence,RI,02912**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**35**	

stochastic control problems in continuous time, for controlled reflected-jump-diffusion type models that have been of interest to date, and converges under broad conditions. For the control problem there are good algorithms for solving the numerical problems, if the dimension is not too high [15]. The method was extended to zero-sum stochastic differential games in [12, 13, 14], with the last two references concerned with the ergodic cost case, extending partial prior results such as [1, 17, 18]. There has been successful numerical work on non-zero-sum differential games [8, 9], based on the Markov chain approximation method, but there does not seem to be any available theory concerning convergence. Works such as [3] are concerned with approximations to non-zero-sum games in normal form and do not apply to the system models or the type of approximations that appear in our numerical approximations.

We will consider a discounted cost problem for a diffusion model in a hyperrectangle G, with absorption on the boundary. The state space G and the boundary absorption are selected only to simplify the development so that we can concentrate on the issues that are unique to the non-zero-sum case. One can replace the hyperrectangle and boundary absorption by an arbitrary convex polyhedron with boundary reflection, if the reflection directions satisfy the conditions in [15] or in [12]. The hyperrectangular state space is often used for purely numerical reasons, to assure a bounded state space, and then it would be large enough so that it would not interfere with the values for the initial conditions of main interest. We will work with two-player games. Any number of players can be dealt with but we stick to two for notational simplicity. The non-zero-sum game is difficult because, as opposed to the zero-sum case, the players are not strictly competitive and have their own value functions.

The idea of the Markov chain approximation method is to first approximate the controlled diffusion dynamics by a suitable Markov chain on a finite state space with a discretization parameter h, then approximate the cost functions. One solves the game problem for the simpler chain model, and then proves that the value functions associated with equilibrium or ϵ-equilibrium strategies for the chain converge to the value functions associated with equilibrium or ϵ_1-equilibrium strategies for the diffusion model, where $\epsilon_1 \to 0$ as $\epsilon \to 0$. The methods of proof are purely probabilistic, no PDE techniques are required, so no knowledge of whatever PDE's yield the equilibrium values are needed.[1] Such methods have the advantage of providing intuition concerning numerical approximations, they cover the bulk of the problem formulations to date, and they converge under quite general conditions. The essential condition is a natural "local consistency" condition. Getting approximations satisfying this condition is usually straightforward. Many methods are discussed in [15] and all of them are applicable to the game problem of interest here. Furthermore, the numerical approximations are represented as processes which are close to the original, which gives the method intuitive meaning. We are not concerned with algorithms for numerically solving the game for the chain model, only showing

[1] At present there seems to be no information available concerning the PDE's that yield the values.

convergence of the solutions to the desired values as the discretization parameter goes to zero.

In Section 2, the model and the cost functions for the players are defined, the boundary conditions discussed and various background material is given. A "uniform in the controls" discrete-time approximation that will be used in the sequel is also given. The convergence proof depends heavily on the fact that the original diffusion process can be approximated (uniformly in the controls), with various approximations to the controls, and the needed results are developed in Section 3. A particular representation of an ϵ-equilibrium strategy, in terms of a "smooth" conditional probability, depending only on selected samples of the driving Wiener process (and not on the entire Wiener process), is given in Section 4. Various facts concerning the Markov chain approximation are collected in Section 5. The reader is referred to [15] for a fuller treatment. The chain is represented in terms of a driving martingale, and this representation is used to get analogs of the results in Section 3 that use approximations to the chain to show that the probability law of the chain and the costs change little if the control process is approximated in various ways. These results are new and should be more broadly useful in dealing with numerical approximations. Theorem 6.1 in Section 6 shows that an 'approximate" equilibrium (value or strategy) for the diffusion is an "approximate" equilibrium (value or strategy) for the chain for small discretization parameter h. If the ϵ-equilibrium value for the chain is unique for small $\epsilon > 0$, then the convergence proof is complete since an "approximate" equilibrium value for the chain is also one for the diffusion. If the value is not unique then the proof of this last fact is more difficult, and we restrict attention to the case where the diffusion coefficient does not depend on the state. This is done in Theorem 6.2, which is a consequence of Theorem 5.6, which, in turn, applies a strong approximation theorem to show that the discrete time approximation to the diffusion and that for the interpolated chain are very close, uniformly in the controls.

2 The Model

We consider systems of the form, where $x(t) \in I\!\!R^v$, Euclidean v-space,

$$x(t) = x(0) + \int_0^t \sum_{i=1}^{2} b_i(x(s), u_i(s))ds + \int_0^t \sigma(x(s))dw(s), \tag{2.1}$$

where player $i = 1, 2$, has control $u_i(\cdot)$ and cost function

$$W_i(u) = E_x^u \int_0^\tau e^{-\beta t} \sum_i k_i(x(s), u_i(s))ds + E_x^u \sum_i e^{-\beta\tau} g_i(x(\tau)). \tag{2.2}$$

Condition (A2.1) below holds, $\beta > 0$, τ is the first time that the boundary of G is hit (it equals infinity if the boundary is never reached), and $w(\cdot)$ is a standard vector-valued Wiener process. The E_x^u denotes the expectation given the use of

control $u(\cdot) = (u_1(\cdot), u_2(\cdot))$ and initial condition x. Define $b(\cdot) = b_1(\cdot) + b_2(\cdot)$, $k(\cdot) = k_1(\cdot) + k_2(\cdot)$.

A2.1. *The functions $b_i(\cdot)$, and $\sigma(\cdot)$ are bounded and continuous and Lipschitz continuous in x, uniformly in u. The controls $u_i(\cdot)$ for player i take values in U_i, a compact set in some Euclidean space, and the functions $k_i(\cdot)$ and $g_i(\cdot)$ are bounded and continuous.*

A control $u_i(\cdot)$ is said to be in $\mathcal{U}_i$, the set of admissible controls for player i, if it is measurable, non-anticipative with respect to $w(\cdot)$, and U_i-valued. Later we will introduce strategies and admissible relaxed controls. The methods of proof use a weak convergence analysis as in [15], and to the extent possible we use the results of that reference. For S a topological space, let $D[S; 0, \infty)$ denote the S-valued functions on $[0, \infty)$ that are right continuous and have left hand limits, and with the Skorohod topology [5, 15] used. If $S = I\!R^v$, then we write $D[S; 0, \infty) = D^v[0, \infty)$.

The first hitting time $\boldsymbol{\tau}$. Getting numerical solutions requires working in a bounded state space. Often the physics of the problem provide both a bounded state space and the proper boundary conditions. Otherwise, "numerical" boundaries are added. In any case, one needs to provide the necessary boundary conditions. These will be equivalent to either reflection or absorption at the boundary. Both are covered in [15]. Here, we chose boundary absorption, but the details that are unique to the non-zero-sum game problem would be the same in both cases.

The nature of the hitting time τ of the boundary of the set G poses a particular concern from the point of view of the convergence of the numerical algorithm. The proof of convergence generates a sequence of process approximations (continuous-time interpolations of the approximating chain) and the exit or boundary hitting times of this sequence has to converge, in an appropriate probabilistic sense, to the exit time of (2.1). In fact, no matter what the numerical procedure, something analogous must take place. In order to see the problem, refer to Figure 1.
In the figure, the sequence of functions $\phi_n(\cdot)$ converges to the limit function $\phi_0(\cdot)$, but the sequence of first contact times (τ_n) of $\phi_n(\cdot)$ converges to a time τ_0 which is not the moment τ of first contact of $\phi_0(\cdot)$ with the boundary line ∂G of G. The problem in this case is that the limit function $\phi_0(\cdot)$ is tangent to ∂G at the time of first contact.

For our control problem, if the approximating costs are to converge to the costs for (2.1), (2.2), then we need to assure (at least with probability one) that the paths of the limit $x(\cdot)$ are not "tangent" to ∂G at the moment τ of first hitting the boundary. For $\phi(\cdot)$ in $D^v[0, \infty)$ (with the Skorokhod topology used), define the function $\hat{\tau}(\phi)$ with values in the compactified infinite interval $\overline{I\!R}^+ = [0, \infty]$ by: $\hat{\tau}(\phi) = \infty$, if $\phi(t) \in G^0$, the interior of G, for all $t < \infty$, and otherwise use

$$\hat{\tau}(\phi) = \inf\{t : \phi(t) \notin G^0\}.$$

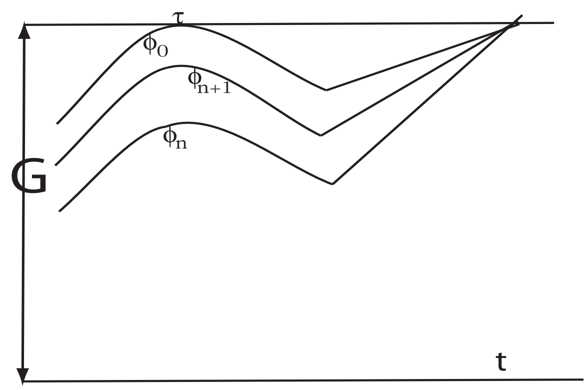

Figure 1: Continuity of first exit times.

In the example of Figure 1, $\hat{\tau}(\cdot)$ is not continuous at the path $\phi_0(\cdot)$.

If the $\phi_0(\cdot)$ in the figure were a sample path of a Wiener process $w(\cdot)$, then the probability is zero that it would be "tangent" to the boundary of G at the point of first contact. Indeed, w.p.1, the path would cross the line infinitely often in any small time interval after first contact. Hence, w.p.1, the first hitting times of any approximating sequence would have to converge to the hitting time of $w(\cdot)$. The situation is similar if the Wiener process were replaced by the solution to a stochastic differential equation with a uniformly positive definite covariance matrix $a(x) = \sigma(x)\sigma'(x)$. The following condition will be used. Note that the condition can be assured to hold if the randomized stopping approximation discussed below is used.

A2.2. *For each initial condition and control, the function* $\hat{\tau}(\cdot)$ *is continuous* (*as a map from* $D^v[0,\infty)$ *to the compactified interval* $[0,\infty]$) *with probability one relative to the measure induced by the solution to* (2.1).

The tangency problem would be a concern with any numerical method, since they all depend on some sort of approximation. For example, the convergence theorems for the classical finite difference methods for elliptic and parabolic equations generally use a nondegeneracy condition on $a(x)$ in order to (implicitly) guarantee (A2.2). In fact, one can always add an independent v-dimensional Wiener process with small variance to (2.1), which will assure (A2.2), while changing the costs arbitrarily little.

The verification of (A2.2) for the case where $a(x)$ is degenerate is more complicated, and one needs to work with the particular structure of the individual case. The boundary can often be divided into several pieces, where we are able to treat each piece separately. For example, there might be a segment where a "directional nondegeneracy" of $a(x)$ guarantees the almost sure continuity of the exit times of the paths which exit on that segment, plus a segment where the direction of the drift gives a similar guarantee, plus a segment on which escape

is not possible, and a "remaining" segment. Frequently, the last "complementary" set is a finite set of points or a curve of lower dimension than that of the boundary. Special considerations concerning these points can often resolve the issue there. An important class of such a degenerate example is illustrated in [10, pp. 64-66]. In that two-dimensional example, G is the symmetric square box centered about the origin and the system is ($x = (x_1, x_2)$)

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= u dt + dw, \end{aligned}$$

and the control $u(\cdot)$ is bounded. The above cited "complementary set" is just the two points which are the intersections of the horizontal axis with the boundary, and these points can be taken care of by a test such as that in Theorem 6.1 of [16]. [2]

Randomized stopping. An alternative to (A2.2). The boundaries in control problems are often not fixed precisely. For example, they might be introduced simply to bound the state space. The original control problem might be defined in an unbounded space, but the space is then truncated for numerical reasons. Even if there is a given "target set," it is often not necessary to fix it too precisely. Such considerations give us some freedom to vary the boundary slightly. The "randomized stopping" alternative discussed next exploits these ideas and assures (A2.2). Under randomized stopping, the probability of stopping at time t (if the process has not yet been stopped) goes to unity as $x(t)$ at that time approaches ∂G. This can be formalized as follows [15].

For some small $\epsilon > 0$, let $\bar{\lambda}(\cdot) > 0$ be a continuous function on the set $N_\epsilon(\partial G) \cap G^0$, where $N_\epsilon(\partial G)$ is the $\epsilon-$neighborhood of the boundary and G^0 is the interior of G. Let $\bar{\lambda}(x) \to \infty$ as x converges to ∂G. Then stop $x(\cdot)$ at time t with stopping rate $\bar{\lambda}(x(t))$ and stopping cost $g_i(x(t))$ for player i. Randomized stopping is equivalent to adding an additional (and state dependent) discount factor which is active near the boundary.

Relaxed controls $\boldsymbol{r_i(\cdot)}$. In control theory, when working with problems concerning convergence of sequences or approximations, it is usual to use the so-called relaxed controls in lieu of ordinary controls. They are used for theoretical purposes only, for the purposes of getting approximation and convergence proofs. Suppose that for some filtration $\{\mathcal{F}_t, t < \infty\}$, standard vector-valued $\mathcal{F}_t$-Wiener process $w(\cdot)$ and for $i = 1, 2$, $r_i(\cdot)$ is a measure on the Borel sets of $U_i \times [0, \infty)$ such that $r_i(U_i \times [0, t]) = t$ and the process $r_i(A \times [0, \cdot])$ is measurable and non-anticipative for each Borel set $A \subset U_i$. Then $r_i(\cdot)$ is said to be an *admissible relaxed control* for player i with respect to $w(\cdot)$ [6, 15]. Abusing notation slightly, we use $\mathcal{U}_i$ for the set of admissible relaxed controls as well for the set of admissible ordinary controls $u_i(\cdot)$. If the Wiener process and filtration are obvious or unimportant, we simply say that $r_i(\cdot)$ is an admissible relaxed control

[2] See also [15, p 280, sec ed.] where it is shown that the Girsanov transformation. can play a useful role in the verification of (A2.2).

for player i. For Borel sets $A \subset U_i$, we will write $r_i(A \times [t_0, t_1]) = r_i(A, [t_0, t_1])$, and write $r_i(A, t_1)$ if $t_0 = 0$. Define $U = U_1 \times U_2$ and $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$. Henceforth $\{\mathcal{F}_t\}$ will denote a filtration such that $w(\cdot)$ is an $\mathcal{F}_t$-standard Wiener process and $r(\cdot)$ is admissible, for the $r(\cdot)$ of concern.

For almost all (ω, t) and each Borel $A \subset U_i$, one can define the left derivative[3]

$$r_i'(A, t) = \lim_{\delta \to 0} \frac{r_i(A, t) - r_i(A, t - \delta)}{\delta}.$$

Without loss of generality, we can suppose that the limit exists for all (ω, t). Then for all (ω, t), $r_i'(\cdot, t)$ is a probability measure on the Borel sets of U_i and for any bounded Borel set B in $U_i \times [0, \infty)$,

$$r_i(B) = \int_0^\infty \int_{U_i} I_{\{(\alpha_i, t) \in B\}} r_i'(d\alpha_i, t) dt.$$

An ordinary control $u_i(\cdot)$ can be represented in terms of the relaxed control $r_i(\cdot)$ that is defined by its derivative, which takes the form $r_i'(A, t) = I_A(u_i(t))$, where $I_A(u_i)$ is unity if $u_i \in A$ and is zero otherwise. The weak topology [15] will be used on the space of admissible relaxed controls. Relaxed controls are commonly used in control theory to prove existence and approximation theorems, since any sequence of relaxed controls has a weakly convergent subsequence. The use of relaxed controls does not change the range of values of the cost functions.

Define the "product" relaxed control $r(\cdot)$, by its derivative $r'(\cdot) = r_1'(\cdot, t) \times r_2'(\cdot, t)$. Thus $r(\cdot)$ is a product measure, with marginals $r_i(\cdot), i = 1, 2$. We will usually write $r(\cdot) = (r_1(\cdot), r_2(\cdot))$ without ambiguity. The pair $(w(\cdot), r(\cdot))$ is called an *admissible pair* if each of the $r_i(\cdot)$ is admissible with respect to $w(\cdot)$. In relaxed control terminology, (2.1) and (2.2) are written as

$$x(t) = x(0) + \sum_{i=1}^{2} \int_0^t \int_{U_i} b_i(x(s), \alpha_i) r_i'(d\alpha_i, s) ds + \int_0^t \sigma(x(s)) dw(s). \tag{2.3}$$

$$W_i(x, r) = E_x^r \int_0^\tau e^{-\beta t} \int_{U_i} \sum_i k_i(x(s), \alpha_i) r_i'(d\alpha_i, s) ds + \sum_i E_x^r e^{-\beta\tau} g_i(x(\tau)). \tag{2.4}$$

The drift terms can be written as (e.g.) $\int_0^t \int_U b(x(s), \alpha) r'(d\alpha, s) ds$.

A discrete time system. We will also have need for the discrete time form

$$\begin{aligned} x^\Delta(n\Delta + \Delta) = x^\Delta(n\Delta) + \int_{n\Delta}^{n\Delta+\Delta} \int_U b(x^\Delta(n\Delta), \alpha) r'(d\alpha, s) ds \\ + \sigma(x^\Delta(n\Delta))[w(n\Delta + \Delta) - w(\Delta)]. \end{aligned} \tag{2.5}$$

We can define the continuous time interpolation $x^\Delta(\cdot)$ either by $x^\Delta(t) = x^\Delta(n\Delta)$

[3] "Left" is used because we need the derivative to be non-anticipative.

for $t \in [n\Delta, n\Delta + \Delta)$, or as (on the same interval)

$$x^\Delta(t) = x^\Delta(n\Delta) + \int_{n\Delta}^t \int_U b(x^\Delta(n\Delta), \alpha) r'(d\alpha, s) ds + \int_{n\Delta}^t \sigma(x^\Delta(n\Delta)) dw(t), \tag{2.6}$$

where it is assumed that $r(t, \cdot)$ is adapted to $\mathcal{F}_{n\Delta-}$, for $t \in [n\Delta, n\Delta + \Delta)$. The associated cost function $W_i^\Delta(x, r)$ is (2.4) with $x^\Delta(\cdot)$ replacing $x(\cdot)$. Let $r^\Delta(\cdot), r(\cdot)$ be admissible relaxed controls with respect to $w(\cdot)$ with $r^\Delta(\cdot) \to r(\cdot)$ w.p.1 (in the weak topology) and $r^\Delta(\cdot)$ adapted as above. Then, as $\Delta \to 0$, the sequence of solutions $\{x^\Delta(\cdot)\}$ also converges w.p.1, uniformly on any bounded time interval and the limit $(x(\cdot), r(\cdot), w(\cdot))$ solves (2.3). By (A2.2), the first hitting times of the boundary also converge w.p.1. to that of the limit. The costs converge as well. The analogous result holds if the randomized stopping alternative is used.

Randomized vs. relaxed controls. For the discrete time system (2.5) or (2.6), the relaxed control can be approximated by a randomized ordinary control (which relates the relaxed control to randomized strategies), as follows. Let $r(\cdot)$ be a relaxed control that is admissible with respect to $w(\cdot)$. Let $\tilde{u}_{i,n}^\Delta$ be a random variable with the (conditional on $\mathcal{F}_{n\Delta}$) distribution $r_{i,n}^\Delta(\cdot) = E_{n\Delta}\left[r_i\left(\cdot, [n\Delta, n\Delta + \Delta]\right)\right]/\Delta$, where $E_{n\Delta}$ denotes the conditional expectation given $\mathcal{F}_{n\Delta}$. Set $\tilde{u}_n^\Delta = (\tilde{u}_{1,n}^\Delta, \tilde{u}_{2,n}^\Delta)$, define its continuous-time interpolation (with intervals Δ) $\tilde{u}^\Delta(\cdot)$, and define the process $\tilde{x}^\Delta(\cdot)$ by $\tilde{x}^\Delta(0) = x^\Delta(0) = x(0)$ and

$$\tilde{x}^\Delta(n\Delta + \Delta) = \tilde{x}^\Delta(n\Delta) + \Delta b(\tilde{x}^\Delta(n\Delta), \tilde{u}_n^\Delta) + \sigma(\tilde{x}^\Delta(n\Delta)[w(n\Delta + \Delta) - w(n\Delta)]. \tag{2.7}$$

Let $\tilde{x}^\Delta(t)$ denote the continuous time interpolation. Then we have the following result, where the relaxed control $r^\Delta(\cdot)$ that is used for $x^\Delta(\cdot)$ has the derivative $r^{\Delta,\prime}(\cdot) = r_{1,n}^\Delta(\cdot) r_{2,n}^\Delta(\cdot)$ on $[n\Delta, n\Delta + \Delta)$. The theorem implies that in the continuous limit, randomized controls turn into relaxed controls.

Theorem 2.1. *Assume* (A2.1)*. Then for any* $T < \infty$*,*

$$\lim_{\Delta \to 0} \sup_{x(0) \in G} \sup_{r \in \mathcal{U}} E \sup_{t \le T} \left| x^\Delta(t) - x(t) \right|^2 = 0, \tag{2.8a}$$

$$\lim_{\Delta \to 0} \sup_{x(0) \in G} \sup_{r \in \mathcal{U}} E \sup_{t \le T} \left| x^\Delta(t) - \tilde{x}^\Delta(t) \right|^2 = 0. \tag{2.8b}$$

Under the additional condition (A2.2) *the costs for* (2.5) *and* (2.7) *converge* (*uniformly in* $x(0), r(\cdot)$) *to those for* (2.3) *as well.*

Comment on the proof. Define $\delta x_n^\Delta = x^\Delta(n\Delta) - \tilde{x}^\Delta(n\Delta)$. Then

$$\begin{aligned}\delta x_{n+1}^\Delta = \delta x_n^\Delta + \Delta \int_U &\left[b(x^\Delta(n\Delta, \alpha) - b(\tilde{x}^\Delta(n\Delta, \alpha)\right] r_n^\Delta(d\alpha) \\ &+ \left[\sigma(x^\Delta(n\Delta)) - \sigma(\tilde{x}^\Delta(n\Delta))\right]\left[w(n\Delta + \Delta) - w(n\Delta)\right] + N_n^\Delta,\end{aligned}$$

where

$$N_n^\Delta = \Delta \left[\int_U b(\tilde{x}^\Delta, \alpha) r_n^\Delta(d\alpha) - b(\tilde{x}^\Delta(n\Delta, \tilde{u}_n^\Delta) \right]$$

is an $\mathcal{F}_{n\Delta}$- martingale difference by the definition of $\tilde{u}_n^\Delta(\cdot)$ via the conditional distribution given $\mathcal{F}_{n\Delta}$. Also $E_{n\Delta}|N_n^\Delta|^2 = O(\Delta^2)$. The proof of the uniform (in the control and initial condition) convergence to zero of $|x^\Delta(\cdot) - \tilde{x}^\Delta(\cdot)|$ and of the differences between the integrals

$$E \int_0^t e^{-\beta t} k(\tilde{x}^\Delta(s), \tilde{u}^\Delta(s)) ds, \quad E \int_0^t \int_U e^{-\beta t} k(x^\Delta(s), \alpha)) r^{\Delta,\prime}(d\alpha, s) ds$$

can then be completed by using the Lipschitz condition and this martingale and conditional variance property. This implies (2.8b). An analogous argument can be used to get (2.8a) for each $r(\cdot)$ and $x(0)$. The facts that (A2.2) holds for (2.3) and that (2.8) holds imply that the stopping times for $x^\Delta(\cdot), \tilde{x}^\Delta(\cdot)$ converge to those for (2.3) as well for each $x(0)$ and $r(\cdot)$.

The uniformity in (2.8a) and in the convergence of the costs can be proved by an argument by contradiction that goes roughly as follows. Suppose, for example, that the uniformity in (2.8a) does not hold. Then take a sequence $x^m(0), r^m(\cdot), \Delta_m \to 0$, $m = 1, 2\ldots$, and associated solutions $x^m(\cdot)$ to (2.3). Let $r_n^{m,\Delta_m}(\cdot)$ be defined as $r_n^\Delta(\cdot)$ was, but based on $r^m(\cdot)$ and let $r^{m,\Delta_m}(\cdot)$ denote the interpolation of the associated relaxed control. Define $x^{m,\Delta_m}(\cdot)$ as the solution to (2.6) with interval Δ_m and controls $r^{m,\Delta_m}(\cdot)$ (alternatively, it could be the piecewise constant interpolation). Suppose that, for some $T < \infty$, $\limsup_{m\to\infty} E \sup_{t\le T} |x^{m,\Delta_m}(t) - \tilde{x}^{m,\Delta_m}(t)|^2 > 0$.

Take an arbitrary weakly convergent subsequence of $x^m(\cdot), x^{m,\Delta_m}(\cdot)$, $r^m(\cdot)$, $r^{m,\Delta_m}(\cdot), w(\cdot)$, with (weak-sense) limit denoted by $x(\cdot), \hat{x}(\cdot), r(\cdot), \hat{r}(\cdot), \hat{w}(\cdot)$. Then it is easy to show that $x(\cdot) = \hat{x}(\cdot)$ and $r(\cdot) = \hat{r}(\cdot)$, that $\hat{w}(\cdot)$ is a standard Wiener process, $\hat{x}(\cdot), \hat{r}(\cdot)$ are non-anticipative with respect to $\hat{w}(\cdot)$ and that the set satisfies (2.3). Assume, without loss of generality, that Skorohod representation is used [5, 15], so that we can suppose that the original and limit processes are all defined on the same probability space and that convergence is w.p.1 in the Skorohod topology. For any $T < \infty$, the set of random variables $\{|x^m(t)|^2, |x^{m,\Delta_m}(t)|^2, t \le T\}$ is uniformly integrable. Thus

$$\lim_{m\to\infty} E \sup_{t\le T} \left| x^{m,\Delta_m}(t) - \hat{x}(t) \right|^2 = 0,$$

and

$$\lim_{m\to\infty} E \sup_{t\le T} |x^m(t) - \hat{x}(t)|^2 = 0,$$

a contradiction to the assertion that the uniformity in $x(0)$ and $r(\cdot)$ in (2.8a) does not hold. ■

3 Classes of Controls and Approximations

The convergence proofs will require the use of special approximations to the general ordinary or relaxed copntrols, and the necessary approximations are

developed in this section and in Theorem 4.1.

For each admissible relaxed control $r(\cdot)$ and $\epsilon > 0$, let $r_i^\epsilon(\cdot)$ be admissible relaxed controls with respect to the same filtration and Wiener process $w(\cdot)$, with derivatives $r_i^{\epsilon,\prime}(\cdot)$, and that satisfy

$$\lim_{\epsilon\to 0} \sup_{r_i\in\mathcal{U}_i} E \sup_{t\le T} \left| \int_0^t \int_{U_i} \phi_i(\alpha_i) \left[r_i'(d\alpha_i, s) - r_i^{\epsilon,\prime}(d\alpha_i, s)\right] ds \right| = 0, \; i = 1, 2, \quad (3.1)$$

for each bounded and continuous real-valued nonrandom function $\phi_i(\cdot)$ and each $T < \infty$. Let $x(\cdot)$ and $x^\epsilon(\cdot)$ denote the solutions to (2.3) corresponding to $r(\cdot)$ and $r^\epsilon(\cdot)$, respectively, with the same $w(\cdot)$ used, but perhaps different initial conditions. In particular, define $x^\epsilon(\cdot)$ by

$$x^\epsilon(t) = x^\epsilon(0) + \int_0^t \int_U b(x^\epsilon(s), \alpha) r^{\epsilon,\prime}(d\alpha, s) ds + \int_0^t \sigma(x^\epsilon(s)) dw(s). \quad (3.2)$$

The processes $x(\cdot)$ and $x^\epsilon(\cdot)$ depend on $r(\cdot)$ and $r^\epsilon(\cdot)$, resp., but this dependence is suppressed in the notation. The next theorem shows that the solution $x(\cdot)$ is continuous in the controls in the sense that (3.3) below holds, and that the costs corresponding to $r(\cdot)$ and $r^\epsilon(\cdot)$ are arbitrarily close for small ϵ, uniformly in $r(\cdot)$.

Theorem 3.1. *Assume* (A2.1). *Let* $(r(\cdot), r^\epsilon(\cdot))$ *satisfy* (3.1) *for each bounded and continuous* $\phi_i(\cdot), i = 1, 2,$ *and* $T < \infty$. *Define* $\delta x^\epsilon(t) = x^\epsilon(t) - x(t)$. *Then for each* t

$$\lim_{\epsilon\to 0} \sup_{x(0), x^\epsilon(0): |x^\epsilon(0) - x(0)| \to 0} \sup_{r\in\mathcal{U}} E\left[\sup_{s\le t} |\delta x^\epsilon(s)| \right]^2 = 0. \quad (3.3)$$

Under the additional condition (A2.2)

$$\lim_{\epsilon\to 0} \sup_{x(0), x^\epsilon(0): |x^\epsilon(0) - x(0)| \to 0} \sup_{r\in\mathcal{U}} |W_i(x, r) - W_i(x, r^\epsilon)| = 0, \quad i = 1, 2. \quad (3.4)$$

Comments on the proof. The proof is very similar to that of Theorem 2.1, and we comment only on the use of (3.1). We can write

$$\begin{aligned} \delta x^\epsilon(t) = \delta x^\epsilon(0) &+ \int_0^t \int_U \left[b(x^\epsilon(s), \alpha) - b(x(s), \alpha)\right] r'(d\alpha, s) ds \\ &+ \int_0^t \left[\sigma(x^\epsilon(s)) - \sigma(x(s))\right] dw(s) \\ &+ \int_0^t \int_U b(x^\epsilon(s), \alpha) \left[r^{\epsilon,\prime}(d\alpha, s) - r'(d\alpha, s)\right] ds \end{aligned} \quad (3.5)$$

It will be seen that the sup over any finite time interval of the absolute value of the last term of (3.5) goes to zero in mean square, by virtue of (3.1). For small

$\lambda > 0$, that term can be rewritten as (modulo $O(\lambda)$)

$$\begin{aligned}
&\sum_{l=0}^{[t/\lambda]-1} \int_{l\lambda}^{l\lambda+\lambda} \int_U b(x^\epsilon(l\lambda),\alpha)\,[r^{\epsilon,\prime}(d\alpha,s) - r'(d\alpha,s)]\,ds \\
&\quad + \sum_{l=0}^{[t/\lambda]-1} \int_{l\lambda}^{l\lambda+\lambda} [b(x^\epsilon(s),\alpha) - b(x^\epsilon(l\lambda),\alpha)]\,[r^{\epsilon,\prime}(d\alpha,s) - r'(d\alpha,s)]\,ds.
\end{aligned} \tag{3.6}$$

Here $[t/\lambda]$ denotes the integer part of t/λ. As $\lambda \to 0$ the expectation of the square of the last term of (3.6) goes to zero, uniformly on any finite time interval, and in $r(\cdot), r^\epsilon(\cdot), x(0), x^\epsilon(0)$, whether or not (3.1) holds, since

$$\lim_{\lambda\to 0} \sup_r \sup_\epsilon E \sup_{l\lambda \le t} \sup_{s\le\lambda} \left|x^\epsilon(l\lambda+s) - x^\epsilon(l\lambda)\right|^2 = 0. \tag{3.7}$$

Assumption (3.1) can be used to show that the same uniform limit in mean square holds for the first term of (3.6) for any λ, as $\epsilon \to 0$. The proof of (3.3) is a consequence of these facts and the Lipschitz condition. The convergence of the costs is a consequence of the convergence of the paths, controls, and an argument concerning the convergence of the stopping times such as used in Theorem 2.1. ■

Finite-valued and piecewise constant approximations $r^\epsilon(\cdot)$ in (3.1). Now some approximations of subsequent interest will be defined. They are just piecewise constant and finite-valued ordinary admissible controls. Consider the following discretization of the U_i. Let $U_i \in I\!R^{d_i}$, Euclidean d_i-space. Given $\mu > 0$, partition $I\!R^{d_i}$ into disjoint (hyper)cubes $\{R_i^{\mu,l}\}$ with diameters μ. The boundaries can be assigned to the subsets in any way. Define $U_i^{\mu,l} = U_i \cap R_i^{\mu,l}$, for the finite number (p_i^μ) of non-empty intersections. Choose a point $\alpha_i^{\mu,l} \in U_i^{\mu,l}$. Now, given admissible $(r_1(\cdot), r_2(\cdot))$, define the approximating admissible relaxed control $r_i^\mu(\cdot)$ on the control value space $U_i^\mu = \{\alpha_i^{\mu,l}, l \le p_i^\mu\}$ by its derivative as $r_i^{\mu,\prime}(\alpha_i^{\mu,l}, t) = r_i'(U_i^{\mu,l}, t)$. Denote the set of such controls by $\mathcal{U}_i(\mu)$. The following theorem is a consequence of Theorem 3.1. A version can also be found in [12].

Theorem 3.2. *Assume* (A2.1)–(A2.2), *and the above approximation of* $r_i(\cdot)$ *by* $r_i^\mu(\cdot) \in \mathcal{U}_i(\mu), i = 1, 2$. *Then* (3.1), (3.3), *and* (3.4) *hold for μ replacing ϵ, no matter what the* $\{U_i^{\mu,l}, \alpha_i^{\mu,l}\}$. *The same result holds if we approximate only one of the* $r_i(\cdot)$.

Finite-valued, piecewise-constant and "delayed" approximations. The proofs of convergence depend on showing that the cost changes little if the control actions of any player are discretized in time and slightly delayed. Let $r_i^\mu(\cdot) \in \mathcal{U}_i(\mu)$, where the control value space for player i is U_i^μ. Let $\Delta > 0$. Define the "backward" differences $\Delta_{i,k}^{\mu,l} = r_i^\mu(\alpha_i^{\mu,l}, k\Delta) - r_i^\mu(\alpha_i^{\mu,l}, k\Delta - \Delta), l \le p_i^\mu$, $k = 1, \ldots$. Define the piecewise constant ordinary controls $u_i^{\mu,\Delta}(\cdot) \in \mathcal{U}_i(\mu)$ on

the interval $[k\Delta, k\Delta+\Delta)$ by

$$u_i^{\mu,\Delta}(t) = \alpha_i^{\mu,l} \text{ for } t \in \left[k\Delta + \sum_{\nu=1}^{l-1} \Delta_{i,k}^{\mu,\nu},\ k\Delta + \sum_{\nu=1}^{l} \Delta_{i,k}^{\mu,\nu}\right). \tag{3.8}$$

Note that on $[k\Delta, k\Delta+\Delta)$, $u_i^{\mu,\Delta}(\cdot)$ takes the value $\alpha_i^{\mu,l}$ on a time interval of length $\Delta_{i,k}^{\mu,l}$. Note also that the $u_i^{\mu,\Delta}(\cdot)$ are "delayed," in that the values of $r_i(\cdot)$ on $[k\Delta-\Delta, k\Delta)$ determine the values of $u_i^{\mu,\Delta}(\cdot)$ on $[k\Delta, k\Delta+\Delta)$. Thus $u_i^{\mu,\Delta}(t), t \in [k\Delta, k\Delta+\Delta)$ is $\mathcal{F}_{k\Delta-}$-measurable. Let $r_i^{\mu,\Delta}(\cdot)$ denote the relaxed control representation of $u_i^{\mu,\Delta}(\cdot)$, with time derivative $r_i^{\mu,\Delta,\prime}(\cdot)$. Let $\mathcal{U}_i(\mu,\delta)$ denote the subset of $\mathcal{U}_i(\mu)$ that are ordinary controls and constant on the intervals $[l\delta, l\delta+\delta), l = 0, 1, \ldots$.

The intervals $\Delta_{i,k}^{\mu,l}$ in (3.8) are just real numbers. For later use, it is important to have them be some multiple of some small $\delta > 0$, where Δ/δ is an integer. Consider one method of doing this. Divide $[k\Delta, k\Delta+\Delta)$ into Δ/δ subintervals of length δ each. Working in order $l = 1, 2\ldots$, to each value $\alpha_i^{\mu,l}$ first assign (the integer part) $[\Delta_{i,k}^{\mu,l}/\delta]$ successive subintervals of length δ. The total fraction of time that is unassigned on any bounded time interval will go to zero as $\delta \to 0$, and how control values are assigned to them will have little effect. However, for specificity for future use consider the following method. The unassigned length for value $\alpha_i^{\mu,l}$ is $L_{i,k}^{\mu,\delta,l} = \Delta_{i,k}^{\mu,l} - [\Delta_{i,k}^{\mu,l}/\delta]\delta$, $i \le p_i^{\mu}$. Define the sum $S_{i,k}^{\mu,\delta} = \sum_l L_{i,k}^{\mu,\delta,l}$, which must be an integral multiple of δ.. Then assign each unassigned δ-interval at random with value $\alpha_{i,k}^{\mu,l}$ chosen with probability $L_{i,k}^{\mu,\delta,l}/S_{i,k}^{\mu,\delta}$. By Theorem 2.1, this assignment and randomization approximates the original relaxed control.

Let $\mathcal{U}_i(\mu,\delta,\Delta)$ denote the set of such controls. If $u_i^{\mu,\delta,\Delta}(\cdot)$ is obtained from $r_i(\cdot)$ in this way, then it is a function of $r_i(\cdot)$, but this functional dependence will be omitted in the notation. Let $r_i^{\mu,\Delta,\delta,\prime}(\cdot)$ denote the time derivative of $r_i^{\mu,\Delta,\delta}(\cdot)$. As stated in the next theorem, which is a consequence of Theorem 3.1, for fixed μ and small δ, $u_i^{\mu,\delta,\Delta}(\cdot)$ well approximates the effects of $u_i^{\mu,\Delta}(\cdot)$ and $r_i(\cdot)$, uniformly in $r_i(\cdot)$ and $\{\alpha_i^{\mu,l}\}$. In particular, (3.1) holds in the sense that for each $\mu > 0$, $\Delta > 0$, and bounded and continuous $\phi_i(\cdot)$, for $i = 1, 2$,

$$\lim_{\delta\to 0} \sup_{r_i \in \mathcal{U}_i} E \sup_{t \le T} \left| \int_0^t \int_{U_i} \phi_i(\alpha_i) \left[r_i^{\mu,\delta,\Delta,\prime}(d\alpha_i, s) - r_i^{\mu,\Delta,\prime}(d\alpha_i, s) \right] ds \right| = 0. \tag{3.9}$$

Theorem 3.3. *Assume* (A2.1)–(A2.2), Let $r_i(\cdot) \in \mathcal{U}_i, i = 1, 2$. *Given* $(\mu, \delta, \Delta) > 0$, *approximate as above the theorem to get* $r_i^{\mu,\delta,\Delta}(\cdot) \in \mathcal{U}_i(\mu,\delta,\Delta)$. *Then* (3.1) *holds for* $r_i^{\mu,\delta,\Delta}(\cdot)$ *and* (μ,δ,Δ) *replacing* $r_i^{\epsilon}(\cdot)$ *and* ϵ, *respectively. Also,* (3.9) *holds. In particular, given* $\epsilon > 0$, *there are* $\mu_\epsilon > 0, \delta_\epsilon > 0, \Delta_\epsilon > 0$ *and* $\kappa_\epsilon > 0$, *such that for* $\mu \le \mu_\epsilon, \delta \le \delta_\epsilon, \Delta \le \Delta_\epsilon$ *and* $\delta/\Delta \le \kappa_\epsilon$,

$$\sup_x \sup_{r_1} \sup_{r_2} \left| W_i(x, r_1, r_2) - W_i(x, r_1, u_2^{\mu,\delta,\Delta}) \right| \le \epsilon. \tag{3.10}$$

The expression (3.10) *holds with the indices* 1 *and* 2 *interchanged or if both controls are approximated.*

Consider the discrete-time system (2.5) *with either the interpolation that is piecewise constant or* (2.6). *Then the* $\mu_\epsilon > 0, \delta_\epsilon > 0, \Delta_\epsilon > 0$ *and* $\kappa_\epsilon > 0$ *can be defined so that*

$$\sup_x \sup_{r_1} \sup_{r_2} \left| W_i(x, r_1, r_2) - W_i^\Delta(x, r_1, u_2^{\mu,\delta,\Delta}) \right| \le \epsilon. \tag{3.11}$$

Define the delayed controls $r_i^\Delta(\cdot)$ *by* $r_i^{\Delta,\prime}(\cdot, s) = r_i'(\cdot, s-\Delta)$ *for* $s \in [n\Delta, n\Delta+\Delta)$. *The* $r_1(\cdot)$ *in* $W_i^\Delta(x, r_1, u_2^{\mu,\delta,\Delta})$ *can be replaced by its approximation* $r_1^{\mu,\Delta}(\cdot)$. *The expression* (3.11) *holds with the indices* 1 *and* 2 *interchanged or if both controls are approximated.*

Note on the initial values of the controls. Since the controls are delayed by Δ, we can assign the values on the initial interval $[0, \Delta]$ in any way at all. Let the values $u_i(l\delta), l\delta \le \Delta$, be in U_i^μ and fixed, for $i = 1, 2$.

4 Equilibria and Approximations

Elliott-Kalton strategies. The classical definition of strategy as used in differential games for models such as (2.1) or (2.3) is that of Elliott and Kalton [4, 7]. A strategy $c_1(\cdot)$ for player 1 is a mapping from $\mathcal{U}_2$ to $\mathcal{U}_1$ with the following property. If admissible controls $r_2(\cdot)$ and $\tilde{r}_2(\cdot)$ satisfy $r_2(s) = \tilde{r}_2(s), s \le t$ for $s \le t$, then $c_1(r_2)(s) = c_1(\tilde{r}_2)(s), s \le t$, and with an analogous definition for player 2 strategies. Let $\mathcal{C}_i$ denote the set of such strategies or mappings for player i. An Elliott-Kalton strategy is a generalization of a feedback control. The current control action that it yields for any player is a function only of the past control actions, and does not otherwise depend on the form of the strategy of the other player.

A pair $\bar{c}_i(\cdot) \in \mathcal{C}_i, i = 1, 2$, is said to be an ϵ-equilibrium strategy pair if for any admissible controls $r_i(\cdot), i = 1, 2,$[4]

$$\begin{aligned} W_1(x, \bar{c}_1, \bar{c}_2) &\ge W_1(x, r_1, \bar{c}_2) - \epsilon, \\ W_2(x, \bar{c}_1, \bar{c}_2) &\ge W_2(x, \bar{c}_1, r_2) - \epsilon. \end{aligned} \tag{4.1}$$

The notation $W_1(x, c_1, c_2)$ implies that each player i uses its strategy $c_i(\cdot)$. When writing $W_i(x, c_1, c_2)$, it is assumed that the associated process is well defined. This will be the case here, since Theorem 3.3 implies that it is sufficient to use strategies whose control functions are piecewise constant. If (4.1) holds with $\epsilon = 0$, then we have an equilibrium strategy pair. The controls can be

[4]The definition in [4] requires that the controls $r_i(A, \cdot)$ be progressively measurable, and not simply measurable and adapted, for each Borel set A. But due to the approximation results of Theorems 3.1–3.3, this added requirement is unnecessary in our case.

either ordinary or relaxed. The notation $W_2(x, c_1, r_2)$ implies that player 1 uses its strategy $c_1(\cdot)$ and player 2 uses the relaxed control $r_2(\cdot)$.

The above definition of strategy does not properly allow for randomized controls, where the realized responses given by the strategy of a player to control process of the other player that are identical on some interval might differ there, depending on the random choices that it makes. So we also allow randomized strategies that have the form of the second line in (4.2) below for either one or both of the players. Theorem 2.1 shows the connection between relaxed and randomized controls, so that one can work with relaxed controls in lieu of randomization, if desired.

We will require the following assumption.

A4.1. For each small $\epsilon > 0$ there is an ϵ-equilibrium Elliott-Kalton strategy $(\bar{c}_1^\epsilon(\cdot), \bar{c}_2^\epsilon(\cdot))$ under which the solution to (2.1) or (2.3) is well defined.

The following approximation theorem will be a key item in the development.

Theorem 4.1. *Assume* (A2.1) *and* (A2.2). *Given $\epsilon_1 > 0$, there are positive numbers μ, δ, Δ, where Δ/δ is an integer, such that the values for any strategy pair $(c_1(\cdot), c_2(\cdot)), i = 1, 2$, with $c_i(\cdot) \in \mathcal{C}_i$ and under which the solution to* (2.3) *is well defined*[5]*, can be approximated within ϵ_1 by strategy pairs $c_i^{\mu,\delta,\Delta}(\cdot), i = 1, 2$, of the following form. The realizations of $c_i^{\mu,\delta,\Delta}(\cdot)$ (which depend on the other player's strategy or control) are ordinary controls in $\mathcal{U}_i(\mu, \delta, \Delta)$, and we denote them by $u_i^{\mu,\delta,\Delta}(\cdot)$. For integer n, k, and $k\delta \in [n\Delta, n\Delta + \Delta)$ and α_i taking values in U_i^μ,*

$$
\begin{aligned}
&P\left\{ u_i^{\mu,\delta,\Delta}(k\delta) = \alpha_i \middle| w(s), s \le k\delta; u_j^{\mu,\delta,\Delta}(l\delta), j = 1, 2, l < k \right\} \\
&= P\left\{ u_i^{\mu,\delta,\Delta}(k\delta) = \alpha_i \middle| w(l\Delta), l \le n; u_j^{\mu,\delta,\Delta}(l\delta), j = 1, 2, l\delta < n\Delta \right\} \\
&= p_{i,k}\left(\alpha_i;\, w(l\Delta), l \le n; u_j^{\mu,\delta,\Delta}(l\delta), j = 1, 2, l\delta < n\Delta\right),
\end{aligned}
\tag{4.2}
$$

which defines the functions $p_{i,k}(\cdot)$. For each positive value of μ, δ, Δ, the functions $p_{i,k}(\cdot)$ can be taken to be continuous in the w-arguments, for each value of the other arguments.

Suppose that the control process realizations for player i are in $\mathcal{U}_i(\mu, \delta, \Delta)$, but those of the other player are general relaxed controls. Then we interpret (4.2)*, applied to that control, as being based on its discretized approximation as derived above Theorem* 3.3.

A convenient representation of the values in (4.2). It will be useful for the convergence proofs if the random selections implied by the conditional probabilties in (4.2) were systematized as follows. Let $\{\theta_l\}$ be random variables that are mutually independent and uniformly distributed on $[0, 1]$. The $\{\theta_k, k \ge l\}$

[5] One or both of them might be simply fixed relaxed feedback controls.

will be independent of all system data before time $l\delta$. For each i, n, k, divide $[0,1]$ into (random) subintervals whose lengths are proportional to the conditional probability of the $\alpha_i^{\mu,l}$ as given by (4.2), and select $\alpha_i^{\mu,l}$ if the random selection on $[0,1]$ falls into that subinterval. The same random variables $\{\theta_l\}$ are used for both players, and for all conditional probability rules of the form (4.2). This representation is used for theoretical purposes only.

Proof. By Theorem 3.3, it is sufficient to work with strategies whose control process realizations are in $\mathcal{U}_i(\mu, \delta, \Delta)$. In any case, let $c_i(\cdot) \in \mathcal{C}_i, i = 1, 2$, be any strategies for which the solution to (2.3) is well defined. Then Theorem 3.3 implies that the control process realizations of the strategies can be approximated by those of a pair of strategies $c_i^{\mu,\delta,\Delta}(\cdot), i = 1, 2$, with control process realizations in $\mathcal{U}_i(\mu, \delta, \Delta), i = 1, 2$. To get the $c_i^{\mu,\delta,\Delta}(\cdot)$, player i would start by calculating the response of $c_i(\cdot)$ to the original strategy of the other player, and then approximate and possibly delay it as done above Theorem 3.3. [I.e., each of the original control sequences is replaced by the discretization discussed above Theorem 3.3.] This approximation is uniform in the original strategies in that the differences in the cost functions can be made small, uniformly in the original strategies, for small enough μ, δ, Δ. Hence Theorem 3.3 yields the claim. The $c_i^{\mu,\delta,\Delta}(\cdot)$ are Elliott-Kalton strategies, since they are simply time and space discretizations of Elliott-Kalton strategies.

The probability law of $(u_1^{\mu,\delta,\Delta}(\cdot), u_2^{\mu,\delta,\Delta}(\cdot), w(\cdot))$ determines the law of the corresponding solution to (2.1). The law of evolution of the controls can be written in recursive form, for $i = 1, 2$, and $k\delta \in [n\Delta, n\Delta + \Delta)$,

$$P\left\{u_i^{\mu,\delta,\Delta}(k\delta) = \alpha_i \middle| w(s), s \le n\Delta; u_j^{\mu,\delta,\Delta}(l\delta), j = 1, 2, l\delta < n\Delta\right\}. \tag{4.3}$$

This yields a "randomized" Elliott-Kalton strategy pair.

Now apply the control rule (4.3) to the piecewise constant interpolation of the discrete-time system (2.5). The probability law of the solution on $[0, t]$ is determined by the law of $\left(u_1^{\mu,\delta,\Delta}(l\delta), u_2^{\mu,\delta,\Delta}(l\delta), l\delta < t; w(n\Delta), n\Delta \le t\right)$. Hence, for $k\delta \in [n\Delta, n\Delta + \Delta)$, the probability law of the controls and paths for $x^\Delta(\cdot)$ can be determined from the formula

$$P\left\{u_i^{\mu,\delta,\Delta}(k\delta) = \alpha_i \middle| w(l\Delta), l \le n; u_j^{\mu,\delta,\Delta}(l\delta), j = 1, 2, l\delta < n\Delta\right\}. \tag{4.4}$$

By Theorem 3.3, for small enough δ, Δ the path $x^\Delta(\cdot)$ is arbitrarily close (uniformly in the original strategies or controls $c_i(\cdot), i = 1, 2$) to the path $x(\cdot)$, both under $u_i^{\mu,\delta,\Delta}(\cdot), i = 1, 2$, where we can suppose (without loss of generality) that the law of evolution of the controls takes the form (4.4). By the same theorem and the construction of the $c_i^{\mu,\delta,\Delta}(\cdot), i = 1, 2$, for small enough μ, δ, Δ this latter path is arbitrarily close (uniformly in the original $c_i(\cdot), i = 1, 2$) to the path $x(\cdot)$ under the original $c_i(\cdot), i = 1, 2$. This argument implies the use of the samples $w(l\Delta)$ in (4.2).

Now turn to the assertion concerning continuity in the w-arguments. (See also [15, Theorem 10.3.1] on this point.) For $\rho > 0$, consider the smoothed conditional probability defined by

$$\begin{aligned} &p^{\rho}_{i,k}\left(\alpha_i;\, w(l\Delta), l \le n; u_j^{\mu,\delta,\Delta}(l\delta), j = 1,2, l\delta < n\Delta\right) \\ &= N(\rho)\int e^{-|z-w|^2/2\rho} p_{i,k}\left(\alpha_i; z; u_j^{\mu,\delta,\Delta}(l\delta), j = 1,2, l\delta < n\Delta\right) dz \end{aligned} \tag{4.5}$$

where $N(\rho)$ is a normalizing constant and $w = \{w(l\Delta), l \le n\}$. The variable z has the same dimension as w. The integral is continuous in the w-variables, uniformly in the others. Also it converges to

$$p_{i,k}\left(\alpha_i;\, w(l\Delta), l \le n; u_j^{\mu,\delta,\Delta}(l\delta), j = 1,2, l\delta < n\Delta\right)$$

for almost all w-values. Hence, by Egoroff's Theorem, it converges almost uniformly in any compact set. For almost all w-values the smoothed conditional probability will choose the same control values as would the original rule defined by (4.3) with a probability that goes to unity as $\rho \to 0$. Hence, without loss of generality we can suppose that the $p_{i,k}(\cdot)$ are smooth in the w-variables, as asserted. ■

5 The Markov Chain Approximation: Brief Review and Approximations

5.1 The Markov Chain Approximation Method

We will start by giving a quick overview of the Markov chain approximation method of [10, 11, 15], starting with some comments for the case where there is only one player. We will then develop some approximation results that are analogous to those in Theorem 3.3, and which will be crucial for the convergence theorems in Section 6. The method consists of two steps. Let $h > 0$ be an approximation parameter. The first step is the determination of a finite-state controlled Markov chain ξ_n^h that has a continuous-time interpolation that is an "approximation" of the process $x(\cdot)$. The second step solves the optimization problem for the chain and a cost function that approximates the one used for $x(\cdot)$. Under a natural "local consistency" condition, the minimal cost function for the chain converges to the minimal cost function for the original problem. In applications, the optimal control for the original problem is also approximated. The approximating chain and local consistency conditions are the same for the game problem. The reference [15] contains a comprehensive discussion of many automatic and simple methods for getting the transition probabilities of the chain. The approximations "stay close" to the physical model and can be adjusted to exploit local features.

The simplest state space for the chain for our model (and the one that we will use for simplicity in the discussion) is based on the regular h-grid S_h in $I\!R^v$.

Define $G_h = S_h \cap G$ and $G_h^0 = S_h \cap G^0$. It is only the points in $G_h^0 \cup \partial G_h$ that are of interest. On G_h^0 the chain "approximates" the diffusion part of (2.1) or (2.3). Let ∂G_h denote the points in $S_h - G_h^0$ that can be reached in one step from G_h^0 under some control. These are the boundary points, and the process stops on first reaching them.

Next we define the basic condition of *local consistency* for the part of a chain ξ_n^h that is on G_h^0. Let $u_n^h = (u_{1,n}^h, u_{2,n}^h)$ denote the controls that are used at step n. Define $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$ and let $E_{x,n}^{h,\alpha}$ denote the expectation given the data to step n (when ξ_n^h has just been computed) with $\xi_n^h = x$ and control value $\alpha = u_n^h$ to be used on the next step. For the game problem, $\alpha = (\alpha_1, \alpha_2)$ with $\alpha_i \in U_i$. Define $a(x) = \sigma(x)\sigma'(x)$. Suppose that there is a function $\Delta t^h(\cdot)$ (this is obtained automatically when the transition probabilities are calculated; see [15] and the example below) such that (this defines the functions $b^h(\cdot)$ and $a^h(\cdot)$)

$$\begin{aligned}
&E_{x,n}^{h,\alpha}\Delta\xi_n^h \equiv b^h(x,\alpha)\Delta t^h(x,\alpha) = b(x,\alpha)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)),\\
&\mathrm{cov}_{x,n}^{h,\alpha}[\Delta\xi_n^h - E_{x,n}^{h,\alpha}\Delta\xi_n^h] \equiv a^h(x,\alpha)\Delta t^h(x,\alpha) = a(x)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)),\\
&\lim_{h\to 0}\sup_{x\in G,\alpha\in U}\Delta t^h(x,\alpha) = 0.
\end{aligned} \tag{5.1}$$

It can be seen that the chain has the "local properties" (conditional mean change and conditional covariance) of the diffusion process.[6] One can always select the transition probabilities such that the intervals $\Delta t^h(x,\alpha)$ do not depend on the control variable, although the general theory in [15] does not require it. Such a simplification is often done in applications to simplify the coding. Let $p^h(x,y|\alpha)$ denote the probability that the next state is y given that the current state is x and control pair $\alpha = (a_1, \alpha_2)$ is used.

Under our condition that the controls are separated in $b(\cdot)$, in that $b(x,\alpha) = b_1(x,\alpha_1) + b_2(x,\alpha_2)$, if desired one can construct the chain so that the controls are "separated" in that the one-step transition probability has the form

$$p^h(x,y|\alpha) = p_1^h(x,y|\alpha_1) + p_2^h(x,y|\alpha_2). \tag{5.2}$$

A useful representation of the transition probabilities. It is useful to have the chains for each h defined on the same probability space, no matter what the controls. This is done as follows. Let $\{\chi_n\}$ be a sequence of mutually independent random variables, uniformly distributed on the interval $[0,1]$ and such that $\{\chi_l, l \geq n\}$ is independent of $\{\xi_l^h, u_l^h, l \leq n\}$. For each value of $x = \xi_n^h, \alpha = u_n^h$, arrange the finite number of possible next possible states y in some order and divide the interval $[0,1]$ into successive subintervals whose lengths are $p^h(x,y|\alpha)$. Then for $x = \xi_n^h, \alpha = u_n^h$, select the next state according to where the (uniformly distributed) random choice for χ_n falls. The same random variables $\{\chi_n\}$ will be used in all cases, for all controls and values of h. This representation is used for theoretical purposes only.

[6]Whether the chain is Markovian or not depends on the form of the control that is applied. But the transition probability will always be locally consistent.

An example of an approximating chain. The simplest case for illustrative purposes is one-dimensional and where h is small enough so that $h|b(\alpha, x)| \le \sigma^2(x)$. Then we can use the transition probabilities and interval, for $x \in G_h^0$ [15, Chapter 5],

$$p^h(x, x \pm h|\alpha) = \frac{\sigma^2(x) \pm hb(x, \alpha)}{2\sigma^2(x)}, \quad \Delta t^h(x, \alpha) = \frac{h^2}{\sigma^2(x)}, \quad \Delta t_n^h = \frac{h^2}{\sigma^2(\xi_n^h)}. \tag{5.3}$$

Admissible controls. Let $\mathcal{F}_n^h$ denote the minimal σ-algebra that measures the control and state data to step n, and let E_n^h denote the expectation conditioned on $\mathcal{F}_n^h$. An admissible control for player i at step n is a U_i-valued random variable that is $\mathcal{F}_n^h$-measurable. Let $\mathcal{U}_i^h$ denote the set of the admissible control processes for player i.

A relaxed control for the chain can be defined as follows. Let $r_{i,n}^h(\cdot)$ be a distribution on the Borel sets of U_i such that $r_{i,n}^h(A)$ is $\mathcal{F}_n^h$-measurable for each Borel set $A \in U_i$. Then the $r_{i,n}^h(\cdot)$ are said to be relaxed controls for player i at step n. As for the model (2.3), an ordinary control at step n can be represented by the relaxed control at step n defined by $r_{i,n}^h(A) = I_{\{u_{i,n}^h \in A\}}$ for each Borel set $A \subset U_i$. Define $r_n^h(\cdot)$ by $r_n^h(A_1 \times A_2) = r_{1,n}^h(A_1) r_{2,n}^h(A_2)$, where the A_i are Borel sets in U_i. The associated transition probability is $\int_U p^h(x, y|a) r_n^h(d\alpha)$. If $r_{i,n}^h(A)$ can be written as a measurable function of ξ_n^h for each Borel set A, then the control is said to be relaxed feedback. Under any feedback (or relaxed feedback or randomized feedback) control, the process ξ_n^h is a Markov chain. More general controls, under which there is more "past" dependence and the chain is not Markovian, will be used as well. Let $\mathcal{C}_i^h$ denote the set of control strategies for ξ_n^h.

The cost function. Discretize the costs as follows. The cost functions are the analogs of (2.2) or (2.4). The cost rate for player i is $k_i(x, \alpha_i)\Delta t^h(x)$. The stopping costs are $g_i(\cdot)$, and τ^h denotes the first time that the set G_h^0 is exited. Let $W_i^h(x, u_1^h, u_2^h)$ denote the expected cost for player i under the control sequences $u_i^h = \{u_{i,n}^h, n \ge 0\}, i = 1, 2$. The numerical problem is to solve the game problem for the approximating chain.

Continuous-time interpolations. The discrete-time chain ξ_n^h is used for the numerical computations. However, for the proofs of convergence, we use a continuous-time interpolation $\xi^h(\cdot)$ of $\{\xi_n^h\}$ that will approximate $x(\cdot)$. This will be a continuous-time process that is constructed as follows. Define $\Delta t_n^h = \Delta t^h(\xi_n^h, u_n^h)$, and $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$. Define $\xi^h(t) = \xi_n^h$ on $[t_n^h, t_{n+1}^h)$. Define the continuous-time interpolations $u_i^h(\cdot)$ of the control actions for player i by $u_i^h(t) = u_{i,n}^h, t_n^h \le t < t_{n+1}^h$, and let its (continuous time) relaxed control representation be denoted by $r_i^h(\cdot)$. Define $r^h(\cdot) = (r_1^h(\cdot), r_2^h(\cdot))$, with time derivative $r^{h,\prime}(\cdot)$. We use $\mathcal{U}_i^h$ for the set of continuous time interpolations of the control for player

i as well. Let τ^h denote the first exit time from G_h^0.

An alternative interpolation. In [15] an interpolation called $\psi^h(\cdot)$ was used as well, and had some advantages in simplifying the proofs there. We describe it briefly so that the convergence results of [15] can be used where needed. For each h, let $\nu_n^h, n = 0, 1, \ldots,$ be mutually independent and exponentially distributed random variables with unit mean, and that are independent of $\{\xi_n^h, u_n^h, n \geq 0\}$. Define $\Delta\tau_n^h = \nu_n^h \Delta t_n^h$, and $\tau_n^h = \sum_{i=0}^{n-1} \Delta\tau_i^h$. Define $\psi^h(t) = \xi_n^h$ and $u_\psi^h(t) = u_n^h$ on $[\tau_n^h, \tau_{n+1}^h)$, Now decompose $\psi^h(\cdot)$ in terms of the continuous-time compensator and martingale. Since the intervals between jumps are $\Delta t_n^h \nu_n^h$, where ν_n^h is exponentially distributed and independent of $\mathcal{F}_n^h$, the jump rate of $\psi^h(\cdot)$ when in state x and under control value α is $1/\Delta t^h(x, \alpha)$. Given a jump, the distribution of the next state is given by the $p^h(x, y|\alpha)$, and the conditional mean change is $b^h(x, \alpha)\Delta t^h(x, \alpha)$. So we can write

$$\psi^h(t) = x(0) + \int_0^t b^h(\psi^h(s), u_\psi^h(s))ds + M^h(t), \tag{5.4}$$

where the martingale $M^h(t)$ has quadratic variation process $\displaystyle\int_0^t a^h(\psi^h(s), u_\psi^h(s))ds$. Under any feedback (or randomized feedback) control, the process $\psi^h(\cdot)$ is a continuous-time Markov chain.

It can be shown that ([15, Sections 5.7.3 and 10.4.1]) there is a martingale $w^h(\cdot)$ (with respect to the filtration generated by the state and control processes, possibly augmented by an "independent" Wiener process) such that

$$M^h(t) = \int_0^t \sigma^h(\psi^h(s), u_\psi^h(s))dw^h(s) = \int_0^t \sigma(\psi^h(s))dw^h(s) + \epsilon^h(t), \tag{5.5}$$

where $\sigma^h(\cdot)[\sigma^h(\cdot)]' = a^h(\cdot)$ (recall the definition of $a^h(\cdot)$ in (5.1)), $w^h(\cdot)$ has quadratic variation It and converges weakly to a standard Wiener process. The martingale $\epsilon^h(\cdot)$ is due to the difference between $\sigma(x)$ and $\sigma^h(x)$ (recall the $o(\Delta t^h)$ terms in (5.1)) and

$$\lim_{h\to 0} \sup_{u^h} \sup_{s\leq t} E|\epsilon^h(s)|^2 = 0 \tag{5.6}$$

for each t. Thus

$$\psi^h(t) = x(0) + \int_0^t \int_U b^h(\psi^h(s), \alpha)r^{h,\prime}(d\alpha, s)ds + \int_0^t \sigma(\psi^h(s))dw^h(s) + \epsilon^h(t). \tag{5.7}$$

The interpolations $\xi^h(\cdot)$ and $\psi^h(\cdot)$ are asymptotically equivalent, as seen in the following theorem, so that any asymptotic results for one are also asymptotic results for the other. We will use $\xi^h(\cdot)$.

Theorem 5.1. *Assume the local consistency* (5.1). *Then the time scales with intervals Δt_n^h and $\Delta\tau_n^h$ are asymptotically equivalent.*

Proof. Let $f^h(t) = \min\{n : t_n^h \geq t\}$. Write $\Delta\tau_n^h - \Delta t_n^h = (\nu_n^h - 1)\Delta t_n^h$, a martingale difference. By the martingale property we have

$$E \sup_{n<f^h(t)} |t_n^h - \tau_n^h|^2 = E \sup_{n<f^h(t)} \left| \sum_{i=0}^{n} \Delta t_i^h (v_n^h - 1) \right| \leq 4E \sum_{i=0}^{f^h(t)} [\Delta t_i^h]^2 E(v_i^h - 1)^2,$$

which goes to zero as $h \to 0$ by the last line of (5.1). The result is the same if we define $f^h(t) = \min\{n : \tau_n^h \geq t\}$. ■

By (5.1), we can write

$$\xi_{n+1}^h = \xi_n^h + b^h(\xi_n^h, u_n^h)\Delta t_n^h + \beta_n^h$$

where β_n^h is a martingale difference with $E_n^h[\beta_n^h][\beta_n^h]' = a^h(\xi_n^h, u_n^h)\Delta t_n^h$. There are martingale differences δw_n^h with conditional (given $\mathcal{F}_n^h$) covariance $\Delta t_n^h I$ such that [15, Section 10.4.1], [10, Section 6.6] $\beta_n^h = \sigma^h(\xi_n^h, u_n^h)\delta w_n^h$. Let $w^h(\cdot)$ denote the continuous time interpolation of $\sum_{i=0}^{n-1} \delta w_n^h$ with intervals Δt_n^h. Then, abusing notation, we can write

$$\begin{aligned} \xi^h(t) = x(0) + \int_0^t b^h(\xi^h(s), u^h(s))ds + \int_0^t \sigma^h(\xi^h(s))dw^h(s) + \epsilon^h(t), \\ \int_0^t \sigma^h(\xi^h(s), u^h(s))dw^h(s) = \int_0^t \sigma(\xi^h(s))dw^h(s) + \epsilon^h(t), \end{aligned} \tag{5.8}$$

where $\epsilon^h(\cdot)$ satisfies (5.6) and is due to the $O(\Delta t^h)$ approximation of $a^h(x, \alpha)$ by $\sigma(x)\sigma(x)'$.

Note on convergence. For any subsequence $h \to 0$, there is a further subsequence (also indexed by h for simplicity) such that $(\xi^h(\cdot), r_1^h(\cdot), r_2^h(\cdot), w^h(\cdot), \tau^h)$ converges weakly to random processes $(x(\cdot), r_1(\cdot), r_2(\cdot), w(\cdot), \tau)$, where $r_i(\cdot)$ is a relaxed control for player i, $(x(\cdot), r_1(\cdot), r_2(\cdot), w(\cdot), I_{\{\tau^h \leq \cdot\}})$ is nonanticipative with respect to the standard vector-valued Wiener process $w(\cdot)$, and, writing $r(\cdot) = (r_1(\cdot), r_2(\cdot))$, the set satisfies

$$x(t) = x(0) + \int_0^t \int_U b(x(s), \alpha) r'(d\alpha, s)ds + \int_0^t \sigma(x(s))dw(s).$$

Also, $W_i^h(x, r_1^h, r_2^h) \to W_i(x, r_1, r_2)$. The proofs of these facts are the same as for the one-player control case in [15, Chapter 10].

On the construction of $\boldsymbol{\delta w^h(\cdot)}$. A special case. Full details for the general method of constructing $w^h(\cdot)$ are in [15, Section 10.4.1], [10, Section 6.6]. To illustrate the idea we will consider a very common, case, and one that will be needed in Theorems 5.2, 5.3, 5.4, 5.6 and 6.2. Suppose that $\sigma(\cdot) = \sigma$ is a constant. Suppose that the components of x can be partitioned as $x = (x^1, x^2)$, and σ can be partitioned as $\sigma = \begin{bmatrix} \sigma^1 & 0 \\ 0 & 0 \end{bmatrix}$ where the dimension of x^1 is d^1,

and σ^1 is a square and invertible matrix of dimension d^1. Partition the $a^h(\cdot)$ in the second line of (5.1) as $a^h(x,\alpha) = \left[\begin{array}{cc} a^{1,h}(x,\alpha) & a^{1,2,h}(x,\alpha) \\ a^{2,1,h}(x,\alpha) & a^{2,h}(x,\alpha) \end{array} \right]$. As $h \to 0$, $a^{1,h}(\cdot) \to \sigma^1[\sigma^1]'$ and all other components go to zero, all uniformly in (x,α). Write the analogous partition $w^h(\cdot) = (w^{1,h}(\cdot), w^{2,h}(\cdot))$. For any Wiener process $w^2(\cdot)$ that is independent of the other random variables, we can let $w^{2,h}(\cdot) = w^2(\cdot)$. The only important component of $w^h(\cdot)$ is $w^{1,h}(\cdot)$ and we can write

$$\begin{aligned}\delta w_n^{1,h} &\equiv w^{1,h}(t_{n+1}^h) - w^{1,h}(t_n^h) \\ &= [\sigma^1]^{-1}\left[\xi_{n+1}^{1,h} - \xi_n^{1,h} - \int_{t_n^h}^{t_{n+1}^h}\int_U b^{1,h}(\xi_n^h,\alpha) r^{h,\prime}(s,d\alpha)ds\right] + \delta\epsilon_n^{1,h},\end{aligned} \tag{5.9}$$

where $\delta\epsilon_n^{1,h}$ is due to the approximation of $\alpha^{1,h}(\cdot)$ by $\sigma^1[\sigma^1]'$ and its interpolation satisfies (5.6). If an ordinary control is used, then the double integral is just $b^1(\xi_n^h, u_n^h)\Delta t_n^h$.

5.2 First approximations to the Chain

Approximation results analogous to those of Theorems 3.1–3.3 can be proved and will be useful. These approximations have an independent interest and should be quite useful for other convergence and approximation analyses for numerical approximations. Theorem 5.2 concerns an approximation to (5.8) that is based on the same $w^h(\cdot)$ process, and will be used in Theorem 6.1. The $w^h(\cdot)$ process depends on the control. For the constant σ-case, Theorem 5.3 shows that this control dependence is small and can be factored out, and (uniform in the control) approximations in terms of an i.i.d. driving sequence are developed. Once this control dependence is factored out, more convenient approximations to the chain can be obtained, and this is done in Theorem 5.4.

Consider the representation (5.8), and for μ, δ, Δ as used in Theorem 3.3 and the $r^h(\cdot) = (r_1^h(\cdot), r_2^h(\cdot))$ in (5.8), define the approximation $u_i^{\mu,\delta,\Delta,h}(\cdot), i = 1,2$, analogously to what was done above Theorem 3.3. For the process $w^h(\cdot)$ that appears in (5.8) under the original control $r^h(\cdot)$, define the process

$$\xi^{\mu,\delta,\Delta,h}(t) = x(0) + \int_0^t b(\xi^{\mu,\delta,\Delta,h}(s), u^{\mu,\delta,\Delta,h}(s))ds + \int_0^t \sigma(\xi^{\mu,\delta,\Delta,h}(s))dw^h(s). \tag{5.10}$$

Let $r_i^{\mu,\delta,\Delta,h}(\cdot)$ denote the relaxed control representation of $u_i^{\mu,\delta,\Delta,h}(\cdot)$. The process defined by (5.10) is not a Markov chain even if the controls are feedback, since the $w^h(\cdot)$ is obtained from the process (5.8) under $r^h(\cdot)$ and not under the $r_i^{\mu,\delta,\Delta,h}(\cdot), i = 1,2$. Let $W_i^{\mu,\delta,\Delta,h}(x, r_1^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h})$ denote the cost for the process (5.10). Define the discrete time system

$$\begin{aligned}\tilde\xi^{\mu,\delta,\Delta,h}(n\Delta+\Delta) = \tilde\xi^{\mu,\delta,\Delta,h}(n\Delta) + \int_0^t b(\tilde\xi^{\mu,\delta,\Delta,h}(n\Delta), u^{\mu,\delta,\Delta,h}(s))ds \\ +\sigma(\tilde\xi^{\mu,\delta,\Delta,h}(n\Delta))[w^h(n\Delta+\Delta) - w^h(n\Delta)],\end{aligned} \tag{5.11}$$

with initial condition $x(0)$ and piecewise-constant continuous-time interpolation denoted by $\tilde{\xi}^{\mu,\delta,\Delta,h}(\cdot)$. Let $\tilde{W}_i^{\mu,\delta,\Delta,h}(x, r_1^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h})$ denote the associated cost. We have the following analog of Theorem 3.3.

Theorem 5.2. *Assume* (A2.1). *Given* $(\mu,\delta,\Delta) > 0$, *approximate* $r_i^h(\cdot)$ *as noted above to get* $r_i^{\mu,\delta,\Delta,h}(\cdot)$. *Given* $\epsilon > 0$, *there are* $\mu_\epsilon > 0, \delta_\epsilon > 0, \Delta_\epsilon > 0$ *and* $\kappa_\epsilon > 0$, *such that for any* $t < \infty$, $\mu \le \mu_\epsilon, \delta \le \delta_\epsilon, \Delta \le \Delta_\epsilon$ *and* $\delta/\Delta \le \kappa_\epsilon$,

$$\lim_{(\mu,\delta,\Delta)\to 0} \sup_{x, r_1^h, r_2^h} E \sup_{s\le t} \left| \xi^{\mu,\delta,\Delta,h}(s) - \xi^h(s) \right| = 0. \tag{5.12}$$

If (A2.2) holds in addition, then

$$\lim_{h\to 0} \sup_{x, r_1^h, r_2^h} \left| W_i^{\mu,\delta,\Delta,h}(x, r_1^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h}) - W_i^h(x, r_1^h, r_2^h) \right| \le \epsilon. \tag{5.13}$$

The expressions (5.12) *and* (5.13) *hold if only one of the controls is approximated, and also if* $\xi^{\mu,\delta,\Delta,h}(\cdot)$ *and* $W_i^{\mu,\delta,\Delta,h}(\cdot)$ *are replaced by* $\tilde{\xi}^{\mu,\delta,\Delta,h}(\cdot)$ *and* $\tilde{W}_i^{\mu,\delta,\Delta,h}(\cdot)$, *resp.*

Comments on the proof. For notational simplicity in the proof drop the superscripts μ, δ. Define $\delta\xi^{\Delta,h}(t) = \tilde{\xi}^{\Delta,h}(t) - \xi^h(t)$. Then, following the procedure of Theorem 3.1, write

$$\begin{aligned}
\delta\xi^{\Delta,h}(t) &= \int_0^t \int_U \left[b(\xi^{\Delta,h}(s),\alpha) - b^h(\xi^h(s),\alpha) \right] r^{h,\prime}(d\alpha, s) ds \\
&+ \int_0^t \left[\sigma(\xi^{\Delta,h}(s)) - \sigma(\xi^h(s)) \right] dw^h(s) \\
&+ \int_0^t \int_U b(\xi^{\Delta,h}(s),\alpha) \left[r^{\Delta,h,\prime}(d\alpha, s) - r^{h,\prime}(d\alpha, s) \right] ds + \epsilon_1(t)
\end{aligned}$$

The $w^h(\cdot), \epsilon_1^h(\cdot)$ are martingales with respect to the filtration induced by the data $(\xi^h(\cdot), r^h(\cdot), w^h(\cdot))$, $w^h(\cdot)$ has quadratic variation It and $\epsilon_1^h(\cdot)$ satisfies (5.6). Partition the last integral analogously to what was done in (3.6), with intervals λ. The process $\xi^{\Delta,h}(\cdot)$ satisfies the following version of (3.7): For any $t > 0$ and small $\kappa > 0$ there is $h_\kappa > 0$ such that for $h \le h_\kappa$,

$$\lim_{\lambda\to 0} \sup_{\mu,\delta,\Delta} \sup_{r^h} E \sup_{l\lambda \le t} \sup_{s\le\lambda} \left| \xi^{\Delta,h}(l\lambda + s) - \xi^{\Delta,h}(l\lambda) \right|^2 \le \kappa.$$

Now, using the martingale property and the Lipschitz condition one proceeds in the same way that would be used for approximations to (2.3) in Theorem 3.1. For example, for some constant K, we have the inequality

$$\begin{aligned}
E \sup_{s\le t} \left| \delta\xi^{\Delta,h}(s) \right|^2 &\le K \int_0^t E \left| \delta\xi^{\Delta,h}(s) \right|^2 ds + \kappa^{\lambda,h}(t) \\
&+ KE \left| \sum_{l=0}^{[t/\lambda]-1} \int_{l\lambda}^{(l+1)\lambda} b(\xi^{\Delta,h}(l\lambda),\alpha) \left[r^{\Delta,h,\prime}(d\alpha, s) - r^{h,\prime}(d\alpha, s) \right] ds \right|^2,
\end{aligned}$$

where $\sup_{s\le t}\kappa^{\lambda,h}(s)\to 0$ as $\lambda\to 0, h\to 0$. For each small λ, the last term in the above expression goes to zero uniformly in h as $(\mu,\delta,\Delta)\to 0$, by the method of approximation of the controls. Then (5.12) follows from the resulting inequality and the Bellman-Gronwall Lemma. The inequality (5.13) follows from (5.12) and (A2.2). ■

5.3 Representations and Approximations of the Chain With Control-Independent Driving Noise

The driving noise $w^h(\cdot)$ depends on the path and control. In Section 6 it will be useful to have approximations to $\xi^h(\cdot)$ (uniform in the control and initial condition) where the driving noise increments are independent of the path and control. To accomplish this we will need to factor $w^h(\cdot)$ as $w^h(\cdot)=\bar{w}^h(\cdot)+\zeta^h(\cdot)$ where $\bar{w}^h(\cdot)$ does not depend on the control and $\zeta^h(\cdot)$ is "asymptotically negligible." We will work with the model described at the end of the first subsection of this section, where $\sigma=\begin{bmatrix}\sigma^1 & 0\\ 0 & 0\end{bmatrix}$, the dimension of x^1 is d^1, and σ^1 is a square and invertible matrix of dimension d^1. The approximation and representation results of Theorems 5.2, 5.3 and 5.5 below will hold for such a form. But to simplify the notation and development, we will work with two specific forms, each of which is typical of a large class of models and numerical algorithms. Case 1 below arises when one uses the so-called central difference approximation. Case 2 arises when one uses a central difference approximation for the non-degenerate part and a one-sided or "upwind" approximation for the degenerate part [15, Chapter 5]. Both forms are locally consistent.

Case 1. Suppose that $d^1=v$, so that σ is invertable. For $a=\sigma\sigma'$, suppose that $a_{i,i}-\sum_{j:j\neq i}|a_{i,j}|\ge 0$. This condition can be weakened if the approximation intervals can depend on the coordinate direction, or if linear transformations of the state space do not pose programming difficulties [15, Chapter 5]. In the same reference it is seen that canonical forms of the transition probabilities and interpolation interval have the form, where $q_{i,j}=q_{j,i}$,

$$\begin{gathered}p^h(x,x\pm e_ih|\alpha)=\frac{q_{i,i}\pm hb_i(x,\alpha)}{Q},\quad \Delta t^h(x,\alpha)=\Delta t^h=\frac{h^2}{Q}\\ p^h(x,x+e_ih+e_jh|\alpha)=p^h(x,x-e_ih-e_jh|\alpha)=\frac{q^+_{i,j}}{Q}\\ p^h(x,x+e_ih-e_jh|\alpha)=p^h(x,x-e_ih+e_jh|\alpha)=\frac{q^-_{i,j}}{Q}\\ Q=2\sum_i q_{i,i}+2\sum_{i,j:i\neq j}|q_{i,j}|\,.\end{gathered}\tag{5.14}$$

The $q_{i,j}$ are defined in terms of the entries of the matrix $\sigma\sigma'$ and are given in [15, Equation (3.15), Chapter 5]. We suppose that h is small enough so that all $q_{i,i}-h|b_i(x,\alpha)|\ge 0$. A simple computation using (5.14) shows that

$b^h(x,\alpha) = b(x,\alpha)$ and $a^h(x,\alpha) = \sigma\sigma' + O(\Delta t^h)$. Also, by (5.14) we can write $\Delta t_n^h = \Delta t^h$. In one dimension, (5.14) reduces to (5.3), where $q_{1,1} = \sigma^2/2$ is determined by the local consistency condition (5.1).

Case 2. Suppose that σ can be partitioned as in the last paragraph of the first subsection of this section: I.e., $\sigma = \begin{bmatrix} \sigma^1 & 0 \\ 0 & 0 \end{bmatrix}$ where the dimension of x^1 is d^1, and σ^1 is a square and invertible matrix of dimension d^1. The problem concerns the effect of the degenerate part. The following canonical model for such cases is motivated by the general model of [15, Chapter 5]. Define $\bar{b} = \sup_{x,\alpha} \sum_{i=d^1+1}^{v} |b_i(x,\alpha)|$. Define $\Delta t^h = \Delta t_n^h = h^2/[Q + h\bar{b}]$. Use the form (5.14) for $i \le d^1$, with Q replaced by $Q + h\bar{b}$. For $i = d^1+1, \ldots, v$, use

$$p^h(x, x \pm e_i h|\alpha) = \frac{h b_i^{\pm}(x,\alpha)}{Q + \bar{b}},$$
$$p^h(x, x|\alpha) = \frac{h\bar{b} - \sum_{i=d^1+1}^{v} |b_i(x,\alpha)|}{Q + h\bar{b}}.$$

We still have $a^h(x,\alpha) = \sigma\sigma' + O(\Delta t^h)$ and $b^h(x,\alpha) = b(x,\alpha)$.

Theorem 5.3. *Use either of the models Case 1 or Case 2 described above. Then we can write $\delta w_n^h = \delta \bar{w}_n^h + \delta\zeta_n^h$, where the components are martingale differences. The $\delta\bar{w}_n^h$ are i.i.d., $\{\delta\bar{w}_l^h, l \ge n\}$ is independent of $\{\xi_l^h, u_l^h, l \le n\}$, and the components have values $O(h)$. Also $E_n^h \delta\bar{w}_n^h [\delta\bar{w}_n^h]' = h^2/Q$,* [7] *and $E_n^h \delta\zeta_n^h[\delta\zeta_n^h]' = O(h\Delta t^h)$, $E_n^h \delta\zeta_n^h [\delta\bar{w}_n^h]' = O(h\Delta t^h)$.*

Proof. The proof is a simple construction. The basic approach is to first define δw_n^h as though $b(\cdot) = 0$. The result will define $\delta\bar{w}_n^h$. Then $d\zeta_n^h$ is defined to make up the difference. The fact that the dominant terms in the transition probabilities in (5.14) do not depend on h, and that the contributions due to the drift (hence control and state) are proportional to h makes this possible. To avoid excessive notation and concentrate on the essential ideas. We start with Case 1 in one dimension. The treatment of the higher-dimensional model follows the same pattern and this is illustrated via a two dimensional case. Then the minor modifications that are required for Case 2 are discussed. The procedure in the general case should be apparent from the three examples.

Case 1, one dimension. Write the double integral term in (5.9) as $b(\xi_n^h, u_n^h)\Delta t^h$, since $b^h(\cdot) = b(\cdot)$. To construct the state transitions, we will use the representation in terms of the random variables χ_n described in the paragraph below (5.2). In one dimension (5.14) is (5.3) and $p^h(x, x \pm h|\alpha) = 0.5 \pm hb(x,\alpha)/[2\sigma^2]$, $\Delta t^h = h^2/\sigma^2$. Now, define $\xi_{n+1}^h - \xi_n^h$ by setting it equal to h if the random sample of χ_n falls in $[0, .5 + hb(\xi_n^h, u_n^h)/2\sigma^2]$, and set it equal to $-h$ otherwise. The "conditional mean" change is $2h[hb(\xi_n^h, u_n^h)/2\sigma^2] = b(\xi_n^h, u_n^h)\Delta t^h$, which is just what is required by the local consistency condition (5.1).

[7] $\Delta t^h = h^2/Q$ for Case 1,

Define the martingale difference term $\delta\bar{w}_n^h$ as follows. Divide $[0,1]$ into the two segments $[0,.5], (.5,1]$, If the random sample of χ_n falls in $[0,.5]$, set $\delta\bar{w}_n^h = h/\sigma$, otherwise set it equal to $-h/\sigma$. It is what δw_n^h would be if $b(\cdot)=0$. Now define $\delta\zeta_n^h$ to make up for the difference. There are two components to $\delta\zeta_n^h$. One component is due to the error $a^h(x,\alpha)-\sigma^2 = O(h^2)$. Hence $[a^h(x,\alpha)]^{1/2}-\sigma = O(h^2)$ and the corresponding error in computing the sample values of δw_n^h is $O(h^3)$. The associated interpolated error process clearly satisfies (5.6).

The second component of $\delta\zeta_n^h$ is due to the neglect of the $b(\cdot)$ in constructing $\delta\bar{w}_n^h$. We handle this as follows. Suppose that $b(\xi_n^h,u_n^h)\ge 0$ (the computation is analogous if $b(\xi_n^h,u_n^h)<0$). Then

$$\delta\zeta_n^h = (2h - b(\xi_n^h,u_n^h)\Delta t^h)/\sigma, \quad \text{if } \chi_n \in [.5, .5+hb(\xi_n^h,u_n^h)/(2\sigma^2])]$$

and it equals $-b(\xi_n^h,u_n^h)\Delta t^h/\sigma$ otherwise. The conditional variance of $\delta\zeta_n^h$ is

$$\begin{aligned} E_n^h\left[2h - b(\xi_n^h,u_n^h)\Delta t^h/\sigma\right]^2 \frac{hb(\xi_n^h,u_n^h)}{2\sigma^2} \\ +E_n^h\left[b(\xi_n^h,u_n^h)\Delta t^h/\sigma\right]^2\left(1-\frac{hb(\xi_n^h,u_n^h)}{2\sigma^2}\right) = O(h)\Delta t^h, \end{aligned}$$

uniformly in the controls. The $\delta\zeta_n^h$ term depends on the control, but the $\delta\bar{w}_n^h$ term does not. It is simply a Bernoulli sequence, with $\{\delta\bar{w}_l^h, l\ge n\}$ independent of the data up to step n. Also, $E_n^h[\delta\bar{w}_n^h]^2 = \Delta t^h$, $E_n^h\delta\bar{w}_n^h\delta\zeta_n^h = O(h)\Delta t^h$ and $E_n^h[\delta\zeta_n^h]^2 = O(h)\Delta t^h$, uniformly in the controls.

Now, construct the continuous-time martingales[8] $\bar{w}^h(t), \zeta^h(t)$ by interpolating the sums $\sum_{i=0}^{n-1}\delta\bar{w}_i^h$ and $\sum_{i=0}^{n-1}\delta\zeta_i^h$ with intervals Δt^h. Write $w^h(t) = \bar{w}^h(t)+\zeta^h(t)$. The $\bar{w}^h(\cdot)$ does not depend on the control, has quadratic variation It, and $\bar{w}^h(s), s\ge t$, is independent of $\xi^h(s), u^h(s), s\le t$. The quadratic variation of $\zeta^h(\cdot)$ (and its quadratic covariation with $\bar{w}^h(\cdot)$) is $O(h)$, uniformly in the controls and initial condition.

Comment on the two-dimensional problem. The following computation illustrates the procedure in higher dimensions. Let $q_{1,2}\ge 0$ for specificity. Divide the unit interval into successive subintervals of lengths $q_{1,1}/Q, q_{1,1}/Q$, $q_{2,2}/Q$, $q_{2.2}/Q, q_{1,2}/Q, q_{1,2}/Q$. Again, the aim is to reproduce the transition probabilities (5.14). If χ_n falls in $[0,(q_{1,1}+hb_1(\xi_n^h,u_n^h))/Q]$, set $\xi_{1,n+1}^h - \xi_{1,n}^h = h$, and $\xi_{2,n+1}^h-\xi_{2,n}^h = 0$. If χ_n falls in $[(q_{1,1}+hb_1(\xi_n^h,u_n^h))/Q, 2q_{1,1}/Q]$, then set $\xi_{1,n+1}^h-\xi_{1,n}^h = -h$, and $\xi_{2,n+1}^h-\xi_{2,n}^h = 0$. Do the analog for the second component, using the two intervals of length $q_{2,2}/Q$. If χ_n falls in the next to last of the four subintervals, then set $\xi_{n+1}^h-\xi_n^h = (h,h)$, and equal to $(-h,-h)$ if χ_n falls in the last of the four subintervals. Define $\delta\bar{w}_n^h$ by repeating the above with $b(x,\alpha)=0$ and premultiplying by σ^{-1}. The procedure is analogous in any dimension.

Comment on Case 2. For ease of presentation, let us work in two dimensions, where only the first component of $x(\cdot)$ has a Wiener process driving term. Then

[8] Actually, martingales when evaluated at the t_n^h, but the difference is unimportant.

$\bar{b} = \max_{x,\alpha} |b_2(x,\alpha)|$ and $Q = 2q_{1,1} = 2[\sigma^1]^2$. Slightly modifying the procedure used for Case 1, divide the unit interval into successive subintervals of lengths

$$\frac{q_{1,1}}{Q+h\bar{b}}, \frac{q_{1,1}}{Q+h\bar{b}}, \frac{h\bar{b}}{Q+h\bar{b}}$$

and divide the last subinterval into two further subintervals of lengths

$$h|b_2(x,\alpha)|/[Q+h\bar{b}], \quad h[\bar{b}-|b_2(x,\alpha)|]/[Q+h\bar{b}].$$

Analogously to what was done in the one-dimensional example of Case 1, if the random sample of χ_n falls in $[0, Q/(Q+h\bar{b})]$ then set $\xi^h_{2,n+1} - \xi^h_{2,n} = 0$. If it falls in the complementary interval $(Q/(Q+h\bar{b}), 1]$, then set $\xi^h_{1,n+1} - \xi^h_{1,n} = 0$. If the random sample of χ_n falls in the last subinterval, then set $\xi^h_{2,n+1} - \xi^h_{2,n} = 0$. If it falls into the next to last subinterval, then set $\xi^h_{2,n+1} - \xi^h_{2,n} = h\text{sign}(b_2(\xi^h_n, u^h_n))$. If it falls into the first (resp., second) subinterval, then set $\xi^h_{1,n+1} - \xi^h_{1,n} = h$ (resp., $-h$). These constructions yield (5.14) with $\Delta t^h = h^2/[Q+h\bar{b}]$.

To get $\delta\bar{w}^h_{1,n}$, repeat the procedure with $b(\cdot) = \bar{b} = 0$ and divide by σ^1. In particular, $\delta\bar{w}^h_{1,n} = h/\sigma^1$ if χ_n falls in $[0, q_{1,1}/Q]$. It is $-h/\sigma^1$ if χ_n falls in $(q_{1,1}/Q, 1]$. The variance is h^2/Q. The value of the second component $\delta\bar{w}^h_{2,n}$ is unimportant since it is eventually multiplied by zero. So, let us use an independent Bernouilli sequence with values $\pm h/\sqrt{Q}$, each taken with probability 1/2.

The terms ζ^h_n for this and the previous example compensate for the errors and is computed using a procedure that is analogous to that in the first Case 1. ■

The theorem implies that $\xi^h(\cdot)$ can be written in the form

$$\xi^h(t) = x(0) + \int_0^t \int_U b(\xi^h(s),\alpha) r^{h,\prime}(d\alpha, s)ds + \int_0^t \sigma(\xi^h(s))d\bar{w}^h(s) + \epsilon^h_2(t), \quad (5.15)$$

where $\epsilon^h_2(\cdot)$ equals $\epsilon^h_1(\cdot)$ plus a stochastic integral with respect to $\zeta^h(\cdot)$, and satisfies (5.6). Since the martingale $\bar{w}^h(\cdot)$ does not depend on the control and is essentially the sum of i.i.d. zero mean random variables of size $O(h)$, the form (5.15) can be used to obtain approximation theorems of the type in Theorems 3.1–3.3. The controls can be space and time discretized with arbitrarily small change in the costs, just as in the cited theorems. For Case 1, the quadratic variation process of $\bar{w}^h(\cdot)$ is It. For Case 2, it is $It[1 + h\bar{b}/Q]$.

Theorem 5.4. *Assume* (A2.1) *and the models of Theorem* 5.3. *Define*

$$\tilde{\xi}^h(t) = x(0) + \int_0^t \int_U b(\tilde{\xi}^h(s),\alpha) r^{h,\prime}(d\alpha, s)ds + \int_0^t \sigma(\tilde{\xi}^h(s))d\bar{w}^h(s). \quad (5.16)$$

Then, for each $t > 0$,

$$\lim_{h\to 0} \sup_{x(0),r^h} E \sup_{s\le t} \left|\xi^h(s) - \tilde{\xi}^h(s)\right|^2 = 0. \quad (5.17)$$

If (A2.2) *is assumed as well, then the costs for the two processes are arbitrarily close, uniformly in the control and initial condition.*

Now, given $(\mu, \delta, \Delta) > 0$*, let* $u_i^{\mu,\delta,\Delta,h}(\cdot)$ *be the delayed and discretized approximation of* $r_i^h(\cdot)$ *that would be defined by the procedure above Theorem* 3.3*, with relaxed control representation of the pair* $(i = 1, 2)$ *of approximations being* $r^{\mu,\delta,\Delta,h}(\cdot)$*. Define the system*

$$\begin{aligned}\xi^{\mu,\delta,\Delta,h}(t) = x(0) + \int_0^t \int_U b(\xi^{\mu,\delta,\Delta,h}(s), \alpha) r^{\mu,\delta,\Delta,h,\prime}(d\alpha, s) ds \\ + \int_0^t \sigma(\xi^{\mu,\delta,\Delta,h}(s)) d\bar{w}^h(s).\end{aligned} \tag{5.18}$$

Then for $t > 0$ *and* $\gamma > 0$ *there are positive numbers* $\mu_\gamma, \delta_\gamma, \Delta_\gamma, h_\gamma, \kappa_\gamma$*, such that for* $\mu \leq \mu_\gamma$*,* $\delta \leq \delta_\gamma$*,* $\Delta \leq \Delta_\gamma$*,* $h \leq h_\gamma$*,* $\delta/\Delta \leq \kappa_\gamma$ *we have*

$$\sup_{r^h, x(0)} E \sup_{s \leq t} \left| \xi^{\mu,\delta,\Delta,h}(s) - \tilde{\xi}^h(s) \right|^2 \leq \gamma. \tag{5.19}$$

If (A2.2) *is assumed as well, then for small* (μ, δ, Δ, h) *the costs are arbitrarily close, uniformly in the control and initial condition.*

Comment on the proof. The proof of the various assertions follows the lines of arguments used in Theorem 5.1, exploiting the martingale properties and the Lipschitz condition. The details are very similar and are omitted.

The terms $[\bar{w}^h(n\Delta+\Delta) - \bar{w}^h(n\Delta)], n = 0, 1, \ldots$ are i.i.d. and have orthogonal components. For Case 1, the covariance Δ is times the identity matrix and for Case 2, it is $\Delta I[1 + h\bar{b}/Q]$, and the processes converge to normally distributed random variables as $h \to 0$. It will be useful to quantify this closeness for use in the next section. This will be done in Theorem 5.6, which requires the following strong approximation theorem for i.i.d. random variables.

Lemma 5.5. [2, Theorem 3.]*Let* $\{\phi_n\}$ *be a sequence of* $I\!R^d$*-valued i.i.d. random variables with zero mean and bounded* $(2 + \delta)$*th moment, where* $0 < \delta \leq 1$*. Suppose that the covariance matrix* Γ *is non-singular. Then without changing the distribution, one can redefine the sequence on a richer probability space together with a Wiener process* $B(\cdot)$ *with covariance matrix* Γ *such that*

$$\left| \sum_{i \leq n} \phi_i - B(n) \right| = o(n^{0.5-c}) \tag{5.20}$$

w.p.1 for large n*, for some* $0 < c < 0.5$.

The following theorem asserts that the interpolated chain can be written essentially as the discrete time system (2.5), which we now write as $x^{\mu,\delta,\Delta}(\cdot)$, when the discretized controls are used.

Theorem 5.6. *Assume* (A2.1) *and* (A2.2) *and the models used in Theorem* 5.3. *Then we can define the probability space such that* $\bar{w}^h(t) = w(t) + \rho^h(t)$, *where* $w(\cdot)$ *is a vector-valued Wiener process with covariance matrix the identity. For each* $t > 0$, $\sup_{s\le t} |\rho^h(s)| \to 0$ *and* $E \sup_{s\le t} |\rho^h(s)|^2 \to 0$ *as* $h \to 0$. *Let* $x^{\mu,\delta,\Delta}(\cdot)$ *be the solution to* (2.5) *with the same Wiener process* $w(\cdot)$ *and with the controls that are used in* (5.18). *Then, for any* $t > 0$,

$$\lim_{h\to 0} \sup_{r^h, x(0)} E \sup_{s\le t} \left| x^{\mu,\delta,\Delta}(s) - \xi^{\mu,\delta,\Delta,h}(s) \right|^2 = 0. \tag{5.21}$$

Proof. Since we have assumed that the same controls are used for both systems (2.5) and (5.18), some explanation is needed. Define ϕ_n by $\delta\bar{w}_n^h = \phi_n \sqrt{h^2/Q}$. This can be done since the parameter h is only a linear scale factor in the construction of the $\delta\bar{w}_n^h$. Then ϕ_n satisfies the conditions of Lemma 5.5 and we can suppose that the probability space is such that (5.20) holds for some Wiener process $B(\cdot)$, whose covariance matrix will be the identity. Then, on this probability space define $\delta\bar{w}_n^h$ in terms of the ϕ_n, as above.

Now, starting with the $\delta\bar{w}_n^h$, one can define (possibly by enlarging the probability space) the chain ξ_n^h and controls u_n^h so that they have the same law as originally. This can be done by using a procedure that is similar to the construction in Theorem 5.2. One starts with the sets in the probability space on which the $\sigma\delta\bar{w}_n^h$ take their particular values. Then modify them by sets whose probabilities are $h|b_i(\xi_n^h, u_n^h)|/Q$ analogously to what was done in Theorem 5.2. One can then construct the chain and controls recursively so that the law of the original process is unchanged. I.e., starting with $x(0)$, get u_0^n which is a (possibly random) function of $x(0)$. Then compute $\delta\bar{w}_0^h$ and then ξ_1^h as just described, and continue. Given the controls u_n^h, they can be time and space discretized and delayed,[9] as in Theorem 5.4.

From Lemma 5.5, we have, w.p.1 for large n,

$$\left| \sum_{i=0}^{n} h\phi_i - hB(t) \right| = ho(n^{0.5-c}). \tag{5.22}$$

Consider Case 1. The process $w(\cdot) = hB(\cdot/\Delta t^h)/\sqrt{Q}$ is a Wiener process whose covariance is the identity. By the above arguments and (5.22), there is a constant $c > 0$ and a $t_h \to 0$ as $h \to 0$ such that for

$$\left| \sum_{i=0}^{t/\Delta t^h} \delta\bar{w}_i^h - w(t) \right| = o(t[\Delta t^h]^c) \tag{5.23}$$

w.p.1 for $t \ge t_h$ and small h. For Case 2, the result is the same since the

[9] Actually, it is only required that the controls be approximated and delayed such that the control applied on $[n\Delta, n\Delta + \Delta)$ is $\mathcal{F}^h_{n\Delta-}$-measurable. The other aspects of the discretization are not needed.

difference between the $\bar{w}^h(\cdot)$ processes for the two cases is

$$\sum_{i=tQ/h^2}^{t(Q+h\bar{b})/h^2} \delta w_i^h,$$

and the sup of this over any finite interval goes to zero in mean square as $h \to 0$.

We can suppose that $w(\cdot)$ is the discrete time process that yields the $w(n\Delta)$ in (2.5). We can also suppose that the controls that were constructed for the chain are applied to (2.5). By the above arguments concerning the approximation of $\bar{w}^h(\cdot)$ by $w(\cdot)$, we can write $\bar{w}^h(t) = w(t) + \rho^h(t)$, where the process $\rho^h(\cdot)$ has independent increments, and $\lim_{h\to 0} E \sup_{s\leq t} |\rho^h(s)|^2 = 0$. From this point on the proof is standard, using the Lipschitz condition and the martingale properties. ■

6 An Approximate Equilibrium for the Diffusion Process is an Approximate Equilibrium for the Chain and Vice Versa

Representations of the transition probability and controls. In the next two theorems, we will use the representations of the transitions of the Markov chain in terms of the i.i.d. random variables $\{\chi_n\}$ discussed in the paragraph after (5.1), and the similar representation for the realizations of the rule (4.2) in terms of the random variables $\{\theta_l\}$ noted in the discussion just below the statement of Theorem 4.1. This assures that the sample path of the approximating chain depends only on the selected control values, and that the selected control value in (4.2) depends only on the past values of the control and Wiener process.

Theorem 6.1. *Assume* (A2.1), (A2.2), *and* (A4.1). *An ϵ-equilibrium value for* (2.1) *or* (2.3) *is an ϵ_1-equilibrium value for the approximating Markov chain, where $\epsilon_1 \to 0$ as $\epsilon \to 0$.*

Proof. Let $\epsilon > 0$ be given. By (A4.1), there is an ϵ-equilibrium strategy pair for (2.3) under which the solution to (2.3) is well defined. By Theorem 4.1, without loss of generality, and for small enough μ, δ and Δ, it can be represented as in (4.2), where we can suppose that Δ/δ is an integer, and the $p_{i,k}(\cdot)$ are continuous in the w variables. We can suppose, w.l.o.g., that for each n, k and i, the rule (4.2) is defined for all possible conditioning u-sequences with values in $U_i^\mu, i = 1, 2$. Let $\bar{c}_1^\Delta(\cdot), \bar{c}_2^\Delta(\cdot)$ denote this strategy pair. The strategies $\bar{c}_i^\Delta(\cdot)$ depend on μ and δ as well as on Δ, but for simplicity we suppress that in the notation. Recall that when a strategy that is defined by a rule such as (4.2) is applied to an arbitrary relaxed control, the formula (4.2) is actually applied to the space-time discretization of that relaxed control, as defined above Theorem

3.3. These strategies $\bar{c}_i^{\Delta}(\cdot)$ will need to be adapted for use on the chain. To do this, simply replace the $w(\cdot)$-samples in (4.2) by samples of the $w^h(\cdot)$ process that was defined in the last section. Keep in mind that these strategies are used for theoretical purposes only, to prove a convergence theorem. They are not for practical implementation. For each integer k, the control value $u_i^{\mu,\delta,\Delta,h}(k\delta)$ that is obtained from the rule (4.2) with $w^h(\cdot)$ used will be applied to the chain for all steps m such that $\tau_m^h \in [k\delta, k\delta+\delta)$. The resulting strategies for the chain will be denoted by $\bar{c}_i^{\Delta,h}(\cdot)$ and are in $\mathcal{C}_i^h$.

We want to show that for small enough (μ, Δ, δ), there are $\epsilon_0 > 0$ and $h_0 > 0$, where $\epsilon_0 \to 0$ as $\epsilon \to 0$ such that for $h \le h_0$ and any sequence $r_i^h(\cdot)$ of admissible relaxed (or ordinary) controls for the chain,

$$\begin{aligned} W_1^h(x, \bar{c}_1^{\Delta,h}, \bar{c}_2^{\Delta,h}) &\ge W_1^h(x, r_1^h, \bar{c}_2^{\Delta,h}) - \epsilon_0, \\ W_2^h(x, \bar{c}_1^{\Delta,h}, \bar{c}_2^{\Delta,h}) &\ge W_2^h(x, \bar{c}_1^{\Delta,h}, r_2^h) - \epsilon_0. \end{aligned} \tag{6.1}$$

The notation $W_2^h(x, \bar{c}_1^{\Delta,h}, r_2^h)$ implies that player 1 uses strategy $\bar{c}_1^{\Delta,h}(\cdot)$ and player 2 uses relaxed control $r_2^h(\cdot)$ (in continuous time interpolation notation), or an ordinary control with this relaxed control representation, with the analogous interpretation when the indices are reversed. The notation $W_1^h(x, \bar{c}_1^{\Delta,h}, \bar{c}_2^{\Delta,h})$ implies that player i uses strategy $\bar{c}_i^{\Delta,h}(\cdot), i = 1, 2$.

Suppose that the pair $\bar{c}_i^{\Delta,h}(\cdot), i = 1, 2$, is used for the chain. Let $\bar{r}_i^{\mu,\delta,\Delta,h}(\cdot), i = 1, 2$, denote the (continuous time interpolation notation) relaxed control representation of the control actions. Let $\xi^h(\cdot)$, $w^h(\cdot)$, and τ^h denote the corresponding interpolation of the chain, the "pre-Wiener" process, and the exit time, resp. The sequence $(\xi^h(\cdot), \bar{r}_1^{\mu,\delta,\Delta,h}(\cdot), \bar{r}_2^{\mu,\delta,\Delta,h}(\cdot),\ w^h(\cdot), \tau^h)$ is tight. Select a weakly convergent subsequence with limit denoted by $(x(\cdot), r_1(\cdot),\ r_2(\cdot), w(\cdot), \tau)$, where $(x(\cdot), r_1(\cdot),\ r_2(\cdot), I_{\{\tau \le \cdot\}})$ is non-anticipative with respect to the standard vector-valued Wiener process $w(\cdot)$, and the set $(x(\cdot), r_1(\cdot), r_2(\cdot), w(\cdot))$ solves (2.3). The limit τ is the first hitting time of the boundary of G by the limit process $x(\cdot)$. The details concerning the tightness, characterization of the limit processes and boundary hitting times, and that they solve (2.3), are the same as for the control problem in [15, Chapters 10, 11].

Henceforth, when weak convergent sequences are dealt with, when needed for simplicity in the argument we will suppose (without loss of generality) that the Skorohod representation is used so that all processes are defined on the same probability space and the weak convergence is equivalent to convergence with probability one in the appropriate topology [5, Theorem 1.8, Chapter 3].

Under Skorohod representation, the rule (4.2) with the $w^h(\cdot)$-samples used converges w.p.1 to the same rule with the $w(\cdot)$-samples used, due to the convergence $w^h(\cdot) \to w(\cdot)$ and the continuity of the probabilities in (4.2) in the w-variables. Because of this, the limits $r_i(\cdot), i = 1, 2$, are just realizations of the original ϵ-equilibrium strategies $\bar{c}_i^{\Delta}(\cdot), i = 1, 2$. Since the solution to (2.1) or (2.3) is unique for each admissible pair (control, Wiener process), we can conclude that the probability law of any limit set $(x(\cdot), r_1(\cdot), r_2(\cdot), w(\cdot))$ is the same, no matter what the selected convergent subsequence. Hence the original set of

processes (before the subsequence was taken) converges weakly to this (unique in the sense of probability law) limit set, where the control is determined by the rules $\bar{c}_i^{\Delta}(\cdot), i = 1, 2$.

By the weak convergence

$$W_1(x, \bar{c}_1^{\Delta}, \bar{c}_2^{\Delta}) \leftarrow W_1^h(x, \bar{c}_1^{\Delta,h}, \bar{c}_2^{\Delta,h}) \le \max_{r_1 \in \mathcal{U}_1^h} W_1^h(x, r_1, \bar{c}_2^{\Delta,h}) = W_1^h(x, \hat{r}_1^h, \bar{c}_2^{\Delta,h}). \tag{6.2}$$

$$W_2(x, \bar{c}_1^{\Delta}, \bar{c}_2^{\Delta}) \leftarrow W_2^h(x, \bar{c}_1^{\Delta,h}, \bar{c}_2^{\Delta,h}) \le \max_{r_2 \in \mathcal{U}_2^h} W_2^h(x, \bar{c}_1^{\Delta,h}, r_2) = W_1^h(x, \bar{c}_1^{\Delta,h}, \hat{r}_2^h). \tag{6.3}$$

It can be shown by a weak convergence argument working with the chain for any fixed $h > 0$ that the maximizing controls $\hat{r}_i^h(\cdot)$ exist. But we need only work with control process that approximate the maximum values arbitrarily well and we suppose that the $\hat{r}_i^h(\cdot)$ are such controls.

It will be shown that

$$\limsup_{h\to 0} W_1^h(x, \hat{r}_1^h, \bar{c}_2^{\Delta,h}) \le W_1(x, \bar{c}_1^{\Delta}, \bar{c}_2^{\Delta}) + \epsilon + \rho(\mu, \delta, \Delta), \tag{6.4}$$

where $\rho(\mu, \delta, \Delta) \to 0$ as $(\mu, \delta, \Delta) \to 0$, with the analogous result for indices $1, 2$ interchanged. Inequalities (6.2), (6.3), and (6.4) imply that if player 2 uses $\bar{c}_2^{\Delta,h}(\cdot)$, then player 1 cannot do better (asymptotically, as $h \to 0$ and modulo $\rho(\mu, \delta, \Delta) + \epsilon$) than by using $\bar{c}_1^{\Delta,h}(\cdot)$, with the analogous result holding for the other player. This last fact implies the theorem since (μ, δ, Δ) can be made as small as desired.

Now (6.4) will be shown. Let $\{u_1^{\mu,\delta,\Delta,h}(l\delta)\}$ denote the values that are obtained from $\hat{r}_1^h(\cdot)$ by the space and time discretization given above Theorem 3.3, and which are used by the rule $\bar{c}_2^{\Delta,h}(\cdot)$. Let $\{u_2^{\mu,\delta,\Delta,h}(l\delta)\}$ denote the control choices for player 2, based on the rule $\bar{c}_2^{\Delta,h}(\cdot)$ and the control of player. Let $r_i^{\mu,\delta,\Delta,h}(\cdot)$ denote the (continuous time) relaxed control representation of $\{u_i^{\mu,\delta,\Delta,h}(l\delta)\}$. The processes $\xi^h(\cdot)$ and $w^h(\cdot)$ now denote the interpolation of the chain and the pre-Wiener process, resp., under the strategy $\bar{c}_2^{\Delta,h}(\cdot)$ and control $\hat{r}_1^h(\cdot)$. This $w^h(\cdot)$ process will be fixed for each h and used in the rest of the proof.

Define the process $\xi^{\mu,\delta,\Delta,h}(\cdot)$ by (5.12), driven by the $\{u_i^{\mu,\delta,\Delta,h}(l\delta)\}, i = 1, 2$, and $w^h(\cdot)$. Note that $\{u_2^{\mu,\delta,\Delta,h}(l\delta)\}$ is the response of $\bar{c}_2^{\Delta,h}(\cdot)$ to any control of player 1 with discretization $\{u_1^{\mu,\delta,\Delta,h}(l\delta)\}$. By Theorem 5.1, we have, for small h,

$$\left| W_i^h(x, \hat{r}_1^h, \bar{c}_2^{\Delta,h}) - W_i^{\mu,\delta,\Delta,h}(x, r_1^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h}) \right| \le \rho_1(\mu, \delta, \Delta) \tag{6.5}$$

where $\rho_1(\mu, \delta, \Delta)$ can be made arbitrarily small, uniformly in $\hat{r}^h(\cdot)$, as $(\mu, \delta, \Delta, h) \to 0$. Also,

$$W_i^{\mu,\delta,\Delta,h}(x, r_1^{\mu,\delta,\Delta,h}, r_2^{\mu,\delta,\Delta,h}) = W_i^{\mu,\delta,\Delta,h}(x, r_1^{\mu,\delta,\Delta,h}, \bar{c}_2^{\Delta,h}). \tag{6.6}$$

Let τ^h denote the first hitting time of the boundary for $\xi^{\mu,\delta,\Delta,h}(\cdot)$.

The set $(\xi\xi^{\mu,\delta,\Delta,h}(\cdot),\hat{r}_1^h(\cdot),r_1^{\mu,\delta,\Delta,h}(\cdot),r_2^{\mu,\delta,\Delta,h}(\cdot),w^h(\cdot),\tau^h)$ is tight. Extract a weakly convergent subsequence, and index it by h also. Denote the limit of the weakly convergent subsequence by $(x(\cdot),\hat{r}_1(\cdot),r_1^{\mu,\delta,\Delta}(\cdot),\, r_2^{\mu,\delta,\Delta}(\cdot).\ w(\cdot),\tau)$. Then, as was the case in an earlier part of the proof, $(x(\cdot),\hat{r}_1(\cdot),r_1^{\mu,\delta,\Delta}(\cdot),\ r_2^{\mu,\delta,\Delta}(\cdot), w(\cdot),I_{\{\tau\le\cdot\}})$ is non-anticipative with respect to the standard Wiener process $w(\cdot)$, the set $(x(\cdot),r_1^{\mu,\delta,\Delta}(\cdot),r_2^{\mu,\delta,\Delta}(\cdot),w(\cdot))$ satisfies (2.3), and τ is the first hitting time of the boundary. The $r_i^{\mu,\delta,\Delta}(\cdot)$ is just the relaxed control that is defined by the weak sense limit $\{u_i^{\mu,\delta,\Delta}(l\delta)\}$ of $\{u_i^{\mu,\delta,\Delta,h}(l\delta)\}$.

We need to show that the limits $u^{\mu,\delta,\Delta}(l\delta)$ are chosen by the conditional probability law that determines $\bar{c}_2^{\Delta}(\cdot)$. I.e., that (along the selected subsequence)

$$\begin{aligned} &p_{2,k}\left(\alpha_2;\ w^h(l\Delta), l\le n; u_j^{\mu,\delta,\Delta,h}(l\delta), j=1,2, l\delta<n\Delta\right)\\ &\qquad\to p_{2,k}\left(\alpha_2;\ w(l\Delta), l\le n; u_j^{\mu,\delta,\Delta}(l\delta), j=1,2, l\delta<n\Delta\right)\end{aligned}\tag{6.7}$$

for $k\delta\in[n\Delta,n\Delta+\Delta)$. In (6.7), the $w^h(\cdot)$ can be replaced by its limit $w(\cdot)$ due to the continuity in w. Since there are only a finite number of values for the control, for any $t<\infty$ the limit $\{u_1^{\mu,\delta,\Delta}(l\delta),u_2^{\mu,\delta,\Delta}(l\delta), l\delta\le t\}$ will be achieved after a finite number of steps through the convergent subsequence, w.p.1. This implies (6.7). [We will comment further on this point at the end of the proof.] Thus the policy $\bar{c}_2^{\Delta}(\cdot)$ acting on any relaxed control with discretization $\{u_1^{\mu,\delta,\Delta}(l\delta)\}$ will yield the sequence $\{u_2^{\mu,\delta,\Delta}(l\delta)\}$. Thus,

$$W_1^{\mu,\delta,\Delta,h}(x,r_1^{\mu,\delta,\Delta,h}(\cdot),\bar{c}_2^{\Delta,h})\to W_1(x,r_1^{\mu,\delta,\Delta}(\cdot),\bar{c}_2^{\Delta})$$

and by (6.5) and (6.6), mod $\rho_1(\mu,\delta,\Delta)$,

$$W_1^h(x,\hat{r}_1^h,\bar{c}_2^{\Delta,h})\to W_1(x,r_1^{\mu,\delta,\Delta}(\cdot),\bar{c}_2^{\Delta}).$$

We can conclude that

$$\begin{aligned}\lim_{h\to 0} W_1^h(x,\hat{r}_1^h,c_2^{\Delta,h}) &\le W_1(x,r_1^{\mu,\delta,\Delta}(\cdot),\bar{c}_2^{\Delta})+\rho_1(\mu,\delta,\Delta)\\ &\le W_1(x,\bar{c}_1^{\Delta},\bar{c}_2^{\Delta})+\rho_1(\mu,\delta,\Delta)+\epsilon,\end{aligned}\tag{6.8}$$

where the ϵ is due to the fact that $(\bar{c}_1^{\Delta}(\cdot),\bar{c}_2^{\Delta}(\cdot))$ is an ϵ-equilibrium. The arbitrariness of the subsequence implies (6.4). The same argument is used when the indices $1,2$ are reversed.

Finally, let us comment on (6.7). Recall that the discretizations given above Theorem 3.3 use fixed (and asymptotically unimportant) values on the initial interval $[0,\Delta)$, so let us use $u_1^{\mu,\delta,\Delta,h}(l\delta)=u_1(l\delta)$, $u_2^{\mu,\delta,\Delta,h}(l\delta)=u_2(l\delta)$ for fixed $u_i(l\delta)$ and $l\delta<\Delta$. For $k\delta\in[\Delta,2\Delta)$, we have the rule

$$p_{2,k}\left(\alpha_2; w^h(\Delta); u_1(l\delta),u_2(l\delta), l\delta<\Delta\right)\tag{6.9}$$

and the probability of selecting any $\alpha_2\in U_2^{\mu}$ converges as $w^h(\cdot)\to w(\cdot)$. Then the limit in (6.9) must be the law of $u_2^{\mu,\delta,\Delta}(l\Delta),\Delta\le l\delta<2\Delta$. Using the method of selecting the control values in terms of the θ_i that was

recalled above the theorem statement, we can suppose that the convergence $u_2^{\mu,\delta,\Delta,h}(l\delta) \to u_2^{\mu,\delta,\Delta}(l\delta), \Delta \le l\delta < 2\Delta$, occurs in a finite number of steps w.p.1, as $h \to 0$ through the convergent subsequence, with the rule (6.9) used. Next, on $[2\Delta, 2\Delta + \Delta)$, we have the rule

$$p_{2,k}\left(\alpha_2; w^h(l\Delta), l \le 2; u_i(l\delta), l\delta < \Delta; u_i^{\mu,\delta,\Delta,h}(l\delta), \Delta \le l\delta < 2\Delta, i = 1, 2\right).$$

The $u_1^{\mu,\delta,\Delta,h}(l\delta), \Delta \le l\delta < 2\Delta$, can be assumed to converge in a finite number of steps as well, w.p.1. Hence, as above, so do the selected values of $u_2^{\mu,\delta,\Delta,h}(l\delta), \Delta \le l\delta < \Delta + 2\Delta$. Continuing in this way yields the form (6.7). ∎

The converse result.

If the ϵ-equilibrium value for the chain is unique for arbitrarily small ϵ, then the converse result is true; namely that ϵ-equilibrium values for the chain are ϵ_1-equilibrium values for (2.3), where $\epsilon_1 \to 0$ as $\epsilon \to 0$, and we are done, since Theorem 6.1 then implies that the ϵ-equilibrium values for the diffusion are also unique for small ϵ, and that the numerical solutions will converge to the desired value. If the ϵ-equilibrium value for the chain is not unique for arbitrarily small ϵ, then we will show that this "converse" assertion is true for the model used in Theorem 5.2. We are not able to show the converse result when $\sigma(\cdot)$ depends on x.

Theorem 6.2. *Assume* (A2.1) *and* (A2.2) *and the model used in Theorem* 5.2. *Then for any* $\epsilon > 0$ *there is* $\epsilon_1 > 0$ *which goes to zero as* $\epsilon \to 0$ *such that an* ϵ*-equilibrium value for the chain* ξ_n^h *for small* h *is an* ϵ_1*-equilibrium value for* (2.3).

Proof. Theorem 5.4 says that the paths and cost functions for (5.15) (which is $\xi^h(\cdot)$ under an arbitrary control), (5.16) (where the control is as in (5.15) but the driving process is $\bar{w}^h(\cdot)$), and (5.18) (which is (5.16) with discretized controls) are arbitrarily close, uniformly in the controls, for small (μ, δ, Δ, h). Theorem 5.6 gives the same result for (5.18) and $x^{\mu,\delta,\Delta}(\cdot)$, which is (2.5) with discretized controls. Theorem 3.3 implies the same thing for $x^{\mu,\delta,\Delta}(\cdot)$ and (2.3). This yields the result. ∎

References

[1] M. Bardi, M. Falcone, and P. Soravia. Numerical methods for pursuit-evasion games via viscosity solutions. In M. Bardi, T.E.S. Raghavan, and T. Parthasarathy, editors, *Stochastic and Differential Games: Theory and Numerical Methods*. Birkhäuser, Boston, 1998.

[2] I. Berkes and W. Philipp. Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.*, 7:29–54, 1979.

[3] E.Altman, O.Pourtallier, A. Haurie, and F. Moresino. Approximating Nash equilibria in nonzero-sum games. *International Game Theory Review*, 2:155–172, 2000.

[4] R.J. Elliott and N.J. Kalton. *Existence of value in differential games,* Mem. AMS, **126**. Amer. Math. Soc, Providence, RI, 1974.

[5] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.

[6] W.F. Fleming. Generalized solutions in optimal stochastic control. In P.T. Liu, E. Roxin, and R. Sternberg, editors, *Differential Games and Control Theory: III*, pages 147–165. Marcel Dekker, 1977.

[7] W.H. Fleming and P.E. Souganidis. On the existence of value functions for two-player zero-sum differential games. *Indiana Univ. Math. J.*, 38:293–314, 1989.

[8] A.B. Haurie, J.B. Krawczyk, and M. Roche. Monitoring cooperative equilibria in a stochastic differential game. *J. of Optimiz. Theory and Applic.*, 81:73–95, 1994.

[9] A.B. Haurie and F. Moreseno. Computing equilibria in stochastic games of intergenerational equity. In *International Conference on Dynamic Games*. International Society of Dynamic Games, 2004.

[10] H.J. Kushner. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York, 1977.

[11] H.J. Kushner. Numerical methods for stochastic control problems in continuous time. *SIAM J. Control Optim.*, 28:999–1048, 1990.

[12] H.J. Kushner. Numerical methods for stochastic differential games. *SIAM J. Control Optim.*, pages 457–486, 2002.

[13] H.J. Kushner. Numerical approximations for stochastic differential games: The ergodic case. *SIAM J. Control Optim.*, 42:1911–1933, 2004.

[14] H.J. Kushner. Numerical methods for stochastic differential games: The ergodic cost criterion. In S. Jorgensen, M. Quincampoix, and T. Vincent, editors, *Annals of the International Society of Dynamic Games*. 2005. To appear.

[15] H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, Berlin and New York, 1992. Second edition, 2001.

[16] D.W. Stroock and S.R.S. Varadhan. On degenerate elliptic and parabolic operators of second order and their associated diffusions. *Comm. Pure Appl. Math.*, 25:651–713, 1972.

[17] M. Tidball. Undiscounted zero-sum differential games with stopping times. In G.J. Olsder, editor, *New Trends in Dynamic Games and Applications*. Birkhäuser, Boston, 1995.

[18] M. Tidball and R.L.V. González. Zero-sum differential games with stopping times: Some results and about its numerical resolution. In T. Basar and A. Haurie, editors, *Advances in Dynamic Games and Applications*. Birkhäuser, Boston, 1994.